

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:06:47 ; Search time 113.524 Seconds
(without alignments)
1425.134 Million cell updates/sec

Title: US-09-676-249D-2

Perfect score: 2299
Sequence: 1 MKKKIKMKRKLGLVFPPLS.....KOPDKQESDILITDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2299	100.0	451	AAU01859	AAU01859 Mycoplasma
2	2299	100.0	451	AAO15861	AAO15861 Mycoplasma
3	2134	92.8	423	AAO15862	AAO15862 Mutant My
4	2109	91.7	423	AAU01860	AAU01860 Mycoplasma
5	714.5	31.1	453	AAK40856	AAK40856 43kd Tegr
6	706	30.7	457	AAO15867	AAO15867 Mycoplasma
7	694.5	30.2	432	AAK67582	AAK67582 Cancer me
8	655	28.5	448	AAK93343	AAK93343 Gene enco
9	394.5	17.2	428	AAV05332	AAV05332 Inflammat
10	394.5	17.2	428	ABP71709	ABP71709 M16i anti
11	393.5	17.1	428	ADK68531	ADK68531 Mycoplasma
12	390.5	17.0	428	ADK68538	ADK68538 Mycoplasma
13	387.5	16.9	429	AAW2727	AAW2727 Membrane
14	384.5	16.7	404	ADK68536	ADK68536 Mycoplasma
15	381.5	16.6	404	ADK68537	ADK68537 Mycoplasma
16	232.5	10.1	350	ABP25843	ABP25843 Streptoco
17	232.5	10.1	350	ABG66912	ABG66912 Streptoco
18	227	9.9	351	AAV81632	AAV81632 Streptoco
19	225	9.8	363	ADK46171	ADK46171 Streptoco
20	223	9.7	350	ABU01234	ABU01234 S. pneumo
21	217.5	9.5	330	ABG66916	ABG66916 Streptoco
22	217.5	9.5	330	ABG66915	ABG66915 Streptoco
23	215.5	9.4	349	ABP27752	ABP27752 Streptoco
24	214.5	9.3	330	ABG66917	ABG66917 Streptoco
25	207	9.0	350	ABB54779	ABB54779 Lactococc

26	203	8.8	328	2	AAW55066	AAW55066 Streptoco
27	203	8.8	328	5	ABP54560	ABP54560 S. pneumo
28	203	8.8	328	7	ADC45089	ADC45089 S. pneumo
29	203	8.8	374	7	ADC95457	ADC95457 E. faecali
30	199.5	8.7	362	7	ADH88119	ADH88119 Enterococ
31	197	8.6	347	5	ABP30886	ABP30886 Streptoco
32	197	8.6	347	5	ABP29678	ABP29678 Streptoco
33	197	8.6	370	5	ABP25842	ABP25842 Streptoco
34	188.5	8.2	347	5	ABP29943	ABP29943 Streptoco
35	188.5	8.2	347	5	ABG66913	ABG66913 Streptoco
36	188	8.2	366	6	ADB09398	ADB09398 Alloiococ
37	182	7.9	339	2	AAV00049	AAV00049 Enterococ
38	182	7.9	339	5	ABP43268	ABP43268 E. faecali
39	182	7.9	339	6	ABU88296	ABU88296 E. faecali
40	182	7.9	339	6	ABU13547	ABU13547 Enterococ
41	182	7.9	361	2	AAV00048	AAV00048 Enterococ
42	182	7.9	361	5	ABP43267	ABP43267 E. faecali
43	182	7.9	361	6	ABU88295	ABU88295 E. faecali
44	182	7.9	361	6	ABU13546	ABU13546 Enterococ
45	182	7.9	375	7	ADH88120	ADH88120 Enterococ

ALIGNMENTS

RESULT 1	
ID	AAU01859 standard; protein; 451 AA.
XX	AAU01859;
AC	07-SEP-2001 (first entry)
DT	
XX	
DE	Mycoplasma hyopneumoniae MHP3 antigen.
XX	
KW	MHP3; antigen; vaccine; enzootic mycoplasma pneumonia; antibody;
KM	immunoassay; immunotherapy; anti-idiotypic antibody.
XX	
OS	Mycoplasma hyopneumoniae.
XX	
FN	Key
FT	Misc-difference 7 Location/Qualifiers
FT	Misc-difference 7 /note= "Encoded by TGA"
FT	Misc-difference 99 /note= "Encoded by TGA"
FT	Misc-difference 138 /note= "Encoded by TGA"
FT	Misc-difference 152 /note= "Encoded by TGA"
FT	Misc-difference 174 /note= "Encoded by TGA"
FT	Misc-difference 198 /note= "Encoded by TGA"
FT	Misc-difference 246 /note= "Encoded by TGA"
XX	
PN	EP1090995-A2.
XX	
PD	11-APR-2001.
XX	
PE	26-SEP-2000; 2000EP-00308421.
XX	
PR	29-SEP-1999; 99US-0156602P.
XX	
PA	(PRTZ) PRTZER PROD INC.
PI	
XX	
XX	King KM, Madura RA, Rosey EL;
DR	WPI: 2001-309781/33.
XX	
XX	N-PSDB; AAS03285.
PT	New apoprotein antigens encoded by mhp3 gene from Mycoplasma hyopneumoniae useful as a vaccine for treating or preventing diseases

PT caused by Mycoplasma hyopneumoniae.
 XX
 PS Claim 11; Page 17-18; 38pp; English.
 XX
 CC The sequence is Mycoplasma hyopneumoniae Mhp3 antigen. Mhp3 antigen and
 CC its fragments are useful in manufacturing a vaccine for treating or
 CC preventing a disease or disorder in an animal, especially pig, caused by
 CC M. hyopneumoniae infection e.g. enzootic mycoplasma pneumonia. The mhp3-
 CC encoded proteins may be used as immunogens to generate antibodies which
 CC immunospecifically bind such an immunogen. The antibodies generated
 CC against the antigen are useful in diagnostic immunoassays, passive
 CC immunotherapy and generation of anti-idiotypic antibodies. Mhp3 proteins
 CC may also be used in immunoassays, e.g. to detect or measure in a
 CC biological sample from a vaccinated or potentially infected test animal
 CC the presence of antibodies to the antigen, and thus to monitor the immune
 CC response and/or to diagnose infection of the animal
 CC
 XX
 SQ Sequence 451 AA;
 Query Match 100.0%; Score 2299; DB 4; Length 451;
 Best Local Similarity 100.0%; Pred. No. 1,4e-164;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKKKIKNNKFLGLVPLSAIATISAGCWDKETTKEKSDADNONKQITDVSKISGLVNE 60
 DB 1 MKKKIKNNKFLGLVPLSAIATISAGCWDKETTKEKSDADNONKQITDVSKISGLVNE 60
 QY 61 RKSEIMAKADANHFGLNMAIYTAGTVNDNSFNQSSWEAIQOLGALTGEITVSVDST 120
 DB 61 RKSEIMAKADANHFGLNMAIYTAGTVNDNSFNQSSWEAIQOLGALTGEITVSVDST 120
 QY 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 DB 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 QY 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 DB 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 QY 181 IPGRYINLTYKTEBAGWLAGYANASFLAKKPSDPTKRSALVIGGISPATVDFIAGYL 240
 DB 181 IPGRYINLTYKTEBAGWLAGYANASFLAKKPSDPTKRSALVIGGISPATVDFIAGYL 240
 QY 241 AGIKAWNLKNSDKKTKITTDKIEINLGFVDQSTYKERLEQIASKDKPSTLLAAGPLTE 300
 DB 241 AGIKAWNLKNSDKKTKITTDKIEINLGFVDQSTYKERLEQIASKDKPSTLLAAGPLTE 300
 QY 301 IFSDIIANQNDRYLIGVTDQSLVYTKTKNKFPTSIILKNLGYSVFSVLSDLYTKKNSRN 360
 DB 301 IFSDIIANQNDRYLIGVTDQSLVYTKTKNKFPTSIILKNLGYSVFSVLSDLYTKKNSRN 360
 QY 361 LAGFPGKSAATVYLGIDRFVDIADTSLSEGNKKLATSEAKKEPEEKTITIPAEV 420
 DB 361 LAGFPGKSAATVYLGIDRFVDIADTSLSEGNKKLATSEAKKEPEEKTITIPAEV 420
 QY 421 RKTLEIPMPDQPDQKQESLDKLITDINKN 451
 DB 421 RKTLEIPMPDQPDQKQESLDKLITDINKN 451
 RESULT 2
 AAO15861
 ID AAO15861 standard; protein; 451 AA.
 XX
 AC AAO15861;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Mycoplasma hyopneumoniae mhp3 protein.
 XX
 KM Apoptotein antigen; enzootic mycoplasma pneumonia; mhp3; vaccine;
 KM Mycoplasma hyopneumoniae infection.
 XX
 OS Mycoplasma hyopneumoniae.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 7

FT /note= "Encoded by TGA"
 FT Misc-difference 99
 FT /note= "Encoded by TGA"
 FT Misc-difference 138
 FT /note= "Encoded by TGA"
 FT Misc-difference 152
 FT /note= "Encoded by TGA"
 FT Misc-difference 174
 FT /note= "Encoded by TGA"
 FT Misc-difference 198
 FT /note= "Encoded by TGA"
 FT Misc-difference 246
 FT /note= "Encoded by TGA"
 FT
 FT EPI245677-A1.
 PN
 PD 02-OCT-2002.
 PP 30-MAR-2001; 2001EP-00303030.
 PR 30-MAR-2001; 2001EP-00303030.
 PA (Pfizer) Pfizer Prod Inc.
 PI King KW, Madura RA, Rosey EL;
 DR WPI; 2002-742716/81.
 DR N-PSDB; AAL50632.
 XX
 PT Novel apoprotein antigens encoded by Mycoplasma hyopneumoniae for use in
 PT vaccines to prevent and treat diseases caused by infection with
 PT Mycoplasma hyopneumoniae in animals, especially pigs.
 PS Claim 11; Fig 1; 38pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of Mycoplasma
 CC hyopneumoniae mhp3 proteins, the invention also comprises novel
 CC apoprotein antigens encoded by the M. hyopneumoniae mhp3 gene. M.
 CC hyopneumoniae is a bacterial pathogen that causes enzootic mycoplasma
 CC pneumonia in pigs. The mhp3 genes, proteins and apoprotein antigens of
 CC the invention are useful in the manufacture of a vaccine for treating/
 CC preventing a disease or disorder caused by infection with M.
 CC hyopneumoniae in an animal, especially a pig. The present amino acid
 CC sequence represents a Mycoplasma hyopneumoniae mhp3 protein of the
 CC invention
 CC
 SQ Sequence 451 AA;
 Query Match 100.0%; Score 2299; DB 5; Length 451;
 Best Local Similarity 100.0%; Pred. No. 1,4e-164;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKKKIKNNKFLGLVPLSAIATISAGCWDKETTKEKSDADNONKQITDVSKISGLVNE 60
 DB 1 MKKKIKNNKFLGLVPLSAIATISAGCWDKETTKEKSDADNONKQITDVSKISGLVNE 60
 QY 61 RKSEIMAKADANHFGLNMAIYTAGTVNDNSFNQSSWEAIQOLGALTGEITVSVDST 120
 DB 61 RKSEIMAKADANHFGLNMAIYTAGTVNDNSFNQSSWEAIQOLGALTGEITVSVDST 120
 QY 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 DB 121 AELEGGTSSLANTNNKNWVLSGFQHGDAFTRWLKIPEKNKIIILGIDWTDTENV 180
 QY 181 IPGRYINLTYKTEBAGWLAGYANASFLAKKPSDPTKRSALVIGGISPATVDFIAGYL 240
 DB 181 IPGRYINLTYKTEBAGWLAGYANASFLAKKPSDPTKRSALVIGGISPATVDFIAGYL 240
 QY 241 AGIKAWNLKNSDKKTKITTDKIEINLGFVDQSTYKERLEQIASKDKPSTLLAAGPLTE 300
 DB 241 AGIKAWNLKNSDKKTKITTDKIEINLGFVDQSTYKERLEQIASKDKPSTLLAAGPLTE 300
 QY 301 IFSDIIANQNDRYLIGVTDQSLVYTKTKNKFPTSIILKNLGYSVFSVLSDLYTKKNSRN 360

```

Db      301  IFSDIANQDRYLIGVDTDQSLVYTKTKKFFSILKNLGYSVFSLDLYTKKNSRN 360
Qy      361  LAGFERGKSAATYVLGKDRPVDIADTSLSGNDKKLATEAISEAKKEFEKTKTIPAEV 420
Db      361  LAGFERGKSAATYVLGKDRPVDIADTSLSGNDKKLATEAISEAKKEFEKTKTIPAEV 420
Qy      421  RKTLEIPEMPDKOPDKQESLDKLITDINKN 451
Db      421  RKTLEIPEMPDKOPDKQESLDKLITDINKN 451

```

RESULT 3

AA015862

ID AA015862 standard; protein; 423 AA.

AC AA015862;

DT 16-JAN-2003 (first entry)

XX Mutant Mycoplasma hyopneumoniae mhp3 protein.

XX Apoprotein antigen; enzootic mycoplasma pneumonia; mhp3; mutant; mutein;
 KW vaccine; Mycoplasma hyopneumoniae infection.

XX Mycoplasma hyopneumoniae.

OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 70

FT /note= "Encoded by RGT"

FT Misc-difference 422

FT /note= "Wild type Lys replaced by Asn"

FT Misc-difference 423

FT /note= "Wild type Asn replaced by Leu"

EP1245677-A1.

02-OCT-2002.

30-MAR-2001; 2001EP-00303030.

30-MAR-2001; 2001EP-00303030.

(PFIZ) PFIZER PROD INC.

King KM, Madura RA, Rosey EL;

WPI; 2002-742716/81.

N-PSDB; AAL50633.

XX Novel apoprotein antigens encoded by Mycoplasma hyopneumoniae for use in
 PT vaccines to prevent and treat diseases caused by infection with
 PT Mycoplasma hyopneumoniae in animals, especially pigs.

PS Claim 4; Page 20-21; 38pp; English.

XX The invention comprises the amino acid and coding sequences of Mycoplasma
 CC hyopneumoniae mhp3 proteins, the invention also comprises novel
 CC apoprotein antigens encoded by the M. hyopneumoniae mhp3 gene. M.
 CC hyopneumoniae is a bacterial pathogen that causes enzootic mycoplasma
 CC pneumonia in pigs. The mhp3 genes, proteins and apoprotein antigens of
 CC the invention are useful in the manufacture of a vaccine for treating/
 CC preventing a disease or disorder caused by infection with M.
 CC hyopneumoniae in an animal, especially a pig. The present amino acid
 CC sequence represents a mutant version of the Mycoplasma hyopneumoniae mhp3
 CC protein which was modified for expression in E. coli

XX Sequence 423 AA;

Query Match 92.8%; Score 2134; DB 5; Length 423;
 Best Local Similarity 99.8%; Pred. No. 3,4e-152;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      30  WDKETTKEEKSADNONKQITDVSKISGLVNERKSEIMAAKADANKHGLMMAIVTAGTV 89
Db      2  WDKETTKEEKSADNONKQITDVSKISGLVNERSEIIMAAADANKHGLMMAIVTAGTV 61
Qy      90  NDNSFNQSSWEAIIQGLALTGSEIITSVDSTAELEGKSSSLANTNNKVVWLSGFQHDAR 149
Db      62  NDNSFNQSGWEAIIQGLALTGSEIITSVDSTAELEGKSSSLANTNNKVVWLSGFQHDAR 121
Qy      150  TRWLKIPENKQLETKNIIILGIDWDTENVITPTGRYINLTYTEEAGWLAGVANASFLA 209
Db      122  TRWLKIPENKQLETKNIIILGIDWDTENVITPTGRYINLTYTEEAGWLAGVANASFLA 181
Qy      210  KKEPSPDPTKSAIIVIGGISPAVTDPIAGYIAGIKAMNLSNPKTKITTDKIEINLQFD 269
Db      182  KKEPSPDPTKSAIIVIGGISPAVTDPIAGYIAGIKAMNLSNPKTKITTDKIEINLQFD 241
Qy      270  VQDSTYERLEQIASKDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVDTDQSLVYTKTK 329
Db      242  VQDSTYERLEQIASKDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVDTDQSLVYTKTK 301
Qy      330  NKEFTSLKNLGYSVFSLDLYTKKNSRNLAGFERGKSAATYVLGDKRFPVDIADTSL 389
Db      302  NKEFTSLKNLGYSVFSLDLYTKKNSRNLAGFERGKSAATYVLGDKRFPVDIADTSL 361
Qy      390  EGNDDKLATPAISEAKKEFEKTKTIPAEVRKTLLEIPEMPDKOPDKQESLDKLITDIN 449
Db      362  EGNDDKLATPAISEAKKEFEKTKTIPAEVRKTLLEIPEMPDKOPDKQESLDKLITDIN 421

```

RESULT 4

AAU01860

ID AAU01860 standard; protein; 423 AA.

AC AAU01860;

07-SEP-2001 (first entry)

Mycoplasma hyopneumoniae recombinant MHP3 antigen.

XX MHP3; antigen; vaccine; enzootic mycoplasma pneumonia; mutant; antibody;
 KW immunoassay; immunotherapy; anti-idiotypic antibody; mutein.

XX Mycoplasma hyopneumoniae.

FH Key Location/Qualifiers

FT Misc-difference 70

FT /note= "Encoded by RGA"

FT Misc-difference 417..422

FT /note= "Encoded by AATTACCGATAT"

EP1090995-A2.

11-APR-2001.

26-SEP-2000; 2000EP-00308421.

29-SEP-1999; 99US-0156602P.

(PFIZ) PFIZER PROD INC.

King KM, Madura RA, Rosey EL;

WPI; 2001-309781/33.

N-PSDB; AAS03286.

XX New apoprotein antigens encoded by mhp3 gene from Mycoplasma
 PT hyopneumoniae useful as a vaccine for treating or preventing diseases
 PT caused by Mycoplasma hyopneumoniae.

PS Claim 4; Page 19-21; 38pp; English.

The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3

CC antigen lacking the first 28 amino acids (the putative signal sequence).
CC MHP3 antigen and its fragments are useful in manufacturing a vaccine for
CC treating or preventing a disease or disorder in an animal, especially
CC pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma
CC pneumonia. The mhp3-encoded proteins may be used as immunogens to
CC generate antibodies which immunospecifically bind such an immunogen. The
CC antibodies generated against the antigen are useful in anti-idiotypic
CC immunosays, passive immunotherapy and generation of anti-idiotypic
CC antibodies. Mhp3 proteins may also be used in immunosays, e.g. to
CC detect or measure in a biological sample from a vaccinated or potentially
CC infected test animal the presence of antibodies to the antigen, and thus
CC to monitor the immune response and/or to diagnose infection of the animal
XX
SQ Sequence 423 AA;
Query Match 91.7%; Score 2109; DB 4; Length 423;
Best Local Similarity 99.8%; Pred. No. 2.6e-150;
Matches 414; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 30 WDKETTKEEKSADNQNQITDVSKISGLVNERKSEIMAKADANKHFGANNAVYAGGTV 89
DB 2 WDKETTKEEKSADNQNQITDVSKISGLVNERKSEIMAKADANKHFGANNAVYAGGTV 61
QY 90 NDNFNGSSWAIQOLGALTGGEITSVDSSTAELEBKSSSLANTNNKNNVLSGQHGDAF 149
DB 62 NDNFNGSSWAIQOLGALTGGEITSVDSSTAELEBKSSSLANTNNKNNVLSGQHGDAF 121
QY 150 TRWLKIPENKOLFTEKNIIIGIDMTPTENVIPTGRVYINLTYYKTEBAGMLAGVANNASPIA 209
DB 122 TRWLKIPENKOLFTEKNIIIGIDMTPTENVIPTGRVYINLTYYKTEBAGMLAGVANNASPIA 181
QY 210 KKPSPDPTKSAIYIGGISPAVTDFIAGYLAGI KAMNLKNSDKKTKITTDKIEINIGFD 269
DB 182 KKPSPDPTKSAIYIGGISPAVTDFIAGYLAGI KAMNLKNSDKKTKITTDKIEINIGFD 241
QY 270 VQDSTKERLEQIASKDKPSTLLAVAGPLTFISFDIINQNDRLIIGVDTQSLVYTTXK 329
DB 242 VQDSTKERLEQIASKDKPSTLLAVAGPLTFISFDIINQNDRLIIGVDTQSLVYTTXK 301
QY 330 NKFTSILKNLGVSVSLDLYTKKNSRMLAGFEFKSATVYLGIDRFVADIATSL 389
DB 302 NKFTSILKNLGVSVSLDLYTKKNSRMLAGFEFKSATVYLGIDRFVADIATSL 361
QY 390 EGNDKLATEAISEAKKEFEKTKTTPAEVVKTLLEIPMDPKQDPKQOESLIDX 444
DB 362 EGNDKLATEAISEAKKEFEKTKTTPAEVVKTLLEIPMDPKQDPKQOESLIDX 416
RESULT 5
AAR40856 ID AAR40856 standard; protein; 453 AA.
XX
AC AAR40856;
XX
DT 25-MAR-2003 (revised)
DT 07-MAR-1994 (first entry)
XX
DE 43kd regression associated antigen.
XX
KW Regression associated antigen; tumour; immunotherapy;
KW anti-idiotypic antibodies; antibodies; tumour regression.
XX
OS Mycoplasma hyorhinis.
FH Key Location/Qualifiers
FT Misc-difference 80 /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT Misc-difference 124 /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT Misc-difference 165 /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT

FT
FT Misc-difference 344 /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT
XX
XX US242823-A.
XX
XX 07-SEP-1993.
XX
XX 02-OCT-1992; 92US-00956546.
XX
XX 07-MAR-1986; 86US-00837494.
XX 16-SEP-1987; 87US-00097910.
XX 11-DEC-1987; 87US-00131815.
XX 04-JAN-1988; 88US-00138923.
XX 16-MAR-1990; 90US-00474730.
XX
XX (TNGE-) INT GENETIC ENG INC.
XX
XX Fareed GC, Sen A, Ghosh-Daastidar P, Jar-How L;
XX WPI; 1993-295229/37.
XX N-PSDB; AAQ47816.
XX
XX DNA encoding a regression-associated antigen from M. hyorhinis - is used
XX to obtain prods. for diagnosis, localisation and therapy of tumours.
XX
XX Disclosure; Fig 3; 40pp; English.
XX
XX Regression associated antigens (RAA's) are identified in material from
XX neoplastic cells by their immunological reactivity with regression
XX associated antibodies from the serum of patients diagnosed as undergoing
XX regression of a tumour. RAA's can be used for tumour immunotherapy and
XX for producing and purifying antibodies which can be used for tumour
XX diagnosis, localisation and therapy. The antibodies can also be used for
XX the production of anti-idiotypic antibodies which can also be used in
XX immunotherapy. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 453 AA;
Query Match 31.1%; Score 714.5; DB 2; Length 453;
Best Local Similarity 38.2%; Pred. No. 3.5e-45;
Matches 175; Conservative 86; Mismatches 150; Indels 45; Gaps 16;
QY 14 GLVPELSAIAI-ISAGCMDEKTKEEKSADNQNQITDVSKISGLVNERKSEIMAKADA 72
DB 12 GTISTVASVATPVS CG---ETDKGKII-----RIPD-----NSFVADROAEIRKA----- 54
QY 73 NKHFGLMAIYVTRAGTVDNDSFNQSSWEAL----QOLGALTG-GEITSVDSSTAELEBK 127
DB 55 -KNDFEFTVLTITAGGTVDQDSFNQSIWEAVLEHYDQIEKTNLDVRSQETNNQSELGKY 113
QY 128 SSLANTNNKNNVLSGFGHDAFTRWLKIPEH-----KOLFTEKNIIIGIDM---TPTEN 179
DB 114 KNFLNGKNNWILTFGQGGQGFPEFKLQOTDSNGKKYSDLLAEKKVITVAVDMDSKEDKO 173
QY 180 VIPTGRVYINLTYYKTEBAGMLAGVANNASPLAKKPSDPTKSAIYIGGISPAVTDFIAGY 239
DB 174 LKAGHFTSILYKTEBAGFIAGYASKEFLAYKFPNDKAKRTIAPFGGHHGAVTDFIAGF 233
QY 240 LAGIKAMNLKNSDKKTKITTDKIEINIGFVDQSTKERLEQIASKDKPSTLLAVAGPLT 299
DB 234 LAGIAKANNNDNPTAKVITISDNNINIDPGF-ISNDKTFPTINGVNNK--SLIVLPVVSGLT 290
QY 300 EIFSDII--ANQNDRLYIGVDTQSLVYTTKKNKFFTSILKNLGVSVSVSLDLYTKKSN 357
DB 291 SSVVDALKSNKDKTKYLIGVDTQSKIFPPA-TVFFTSIEKHLARTIYEVLTIDWLKED 349
QY 358 SRNLAGEFEFGK---KSATVYLGIKORFVADIATSLBGNDKLATEAISEAKKEFEKTK 413
DB 350 SKFLGSFRSPKLTNPANATYTKGISDPFGVSNSTVADADKVKYKQEFLENTADTFKQIQ 409
QY 414 TTPAEVVKTLLEIPDM--PDKQDPKQOESLIDXITDIN 449

Db 410 ANPT-NYKSVLGIPFTMLINDNDADAKNEKALNELIKIN 446

RESULT 6

AAOI5867 standard; protein; 457 AA.

AAOI5867;

16-JAN-2003 (first entry)

Mycoplasma hyorhinis Ag234-5 protein.

Apoptosein antigen; enzootic mycoplasma pneumoniae; mhp3; Ag234-5; vaccine; Mycoplasma hyopneumoniae infection.

Mycoplasma hyorhinis.

EP1245677-A1.

02-OCT-2002.

30-MAR-2001; 2001EP-00303030.

30-MAR-2001; 2001EP-00303030.

(PRIZ) PRIZER PROD INC.

King KW, Madura RA, Rosey EL;

WPI; 2002-742716/81.

Novel apoptosein antigens encoded by Mycoplasma hyopneumoniae for use in vaccines to prevent and treat diseases caused by infection with Mycoplasma hyopneumoniae in animals, especially pigs.

Example; Fig 1; 38pp; English.

The invention comprises the amino acid and coding sequences of Mycoplasma hyopneumoniae mhp3 proteins, the invention also comprises novel apoptosein antigens encoded by the M. hyopneumoniae mhp3 gene. M. hyopneumoniae is a bacterial pathogen that causes enzootic mycoplasma pneumonia in pigs. The mhp3 genes, proteins and apoptosein antigens of the invention are useful in the manufacture of a vaccine for treating/preventing a disease or disorder caused by infection with M. hyopneumoniae in an animal, especially a pig. The present amino acid sequence represents the Mycoplasma hyorhinis Ag234-5 protein which is shown in a figure of the invention

Sequence 457 AA:

Query Match 30.7%; Score 706; DB 5; Length 457;
Best Local Similarity 38.4%; Pred. No. 1.5e-44;
Matches 173; Conservative 87; Mismatches 145; Indels 46; Gaps 16;

14 GLVPEPLAIAIT-ISAQCMDEKTEKESADNONQITDVSKISGLVNERKSEIMAAKADA 72
12 GTISTVASVATFVSCG---ETDREGKIT---RIFF---NSFVKDRQAEIERA--- 54
73 NKHFGLMAIYTAGTVDNSFNQSSWEAI-----QQLGALTG-GEITVSDSTALBEGKY 127
55 -KNDFPMTVLITAGTVDKSFNQSIMEVALEHVDQIEKTNLDRVSETNNGSELIGKY 113
128 SSLANTKQVAVLSGFQHGDAFTWMLKIPEN-----KOLFTEKNIITIGIDW---TDTEN 179
114 KNPLNGNKVWILITGFOGGEFFPKFLKQTDNSNGKYSDDLAEKVIIVAVMDLSKEDKD 173
180 VIPIGRYNTLYTKTEEGWGLVYANASFLAKKPPSDPKRSIAIYGGISPAVTDFTAGY 239
174 LTKKHGHSLLYKTEEGWGLVYANASFLAKKPPSDPKRSIAIYGGISPAVTDFTAGY 233
240 LAGIKAVNLKNSDKTKITTDKIEINLGFVDVODTSTKERLEQIASKDKPSTLLAVAGPLT 299

Db 234 LAGIAKTNNDNPTAKVTISDNNINIDYGF-ISNDKATPFIINGIVNKS--SLVLPVAGSLT 290
Qy 300 EIFSDII--ANQNRVYIGVDTOGSLVYTKKNEFTSILKNLGVSFVSVDLYTKSN 357
Db 291 SSVVDALKSNKDKTYLIGVDTQSKIFSPA-TVFFSIEKHGLRRTTYQVLTIDWLKED 349
Qy 358 SRNLAGEFPGK---KSATVYLGIDRFVDIADTSLGNDPKKLATEAISEAKKEFEKTK 413
Db 350 SKFLGSFRSPFLTNPAATYKGIISDPFGVNSNTVADADKVKQGEFLNATADFKKIQ 409
Qy 414 TIPAEVRKTLTEIPEW--PKQDPKQESL 441
Db 410 ANPT-NYKSVLGIPFTMLINDNDADAKNEKASL 439

RESULT 7

AA67582 standard; protein; 432 AA.

AA67582;

06-OCT-1995 (first entry)

Cancer metastasis protein.

Cancer metastasis; clinical tests; detection.

Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 55 /note= "corresponding codon TGA"

FT Misc-difference 99 /note= "corresponding codon TGA"

FT Misc-difference 140 /note= "corresponding codon TGA"

FT Misc-difference 319 /note= "corresponding codon TGA"

FT Misc-difference 420 /note= "corresponding codon TGA"

FT Misc-difference 424 /note= "corresponding codon TGA"

PN JP06319559-A.

PD 22-NOV-1994.

PP 12-MAY-1993; 93JP-00144165.

PR 12-MAY-1993; 93JP-00144165.

PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU.

DR WPI; 1995-040317/06.

DR N-PSDB; AAQ79124.

FT DNA encoding polypeptide involved in cancer metastasis - useful for the study of the mechanism of metastasis and in clinical tests.

PS Claim 1; Page 2; 13pp; Japanese.

AAQ79124 encodes AAR67582 a protein involved in cancer metastasis, which may be used for the study of metastatic mechanisms, and for clinical tests to determine the presence or absence of cancer metastasis

Sequence 432 AA:

Query Match 30.2%; Score 694.5; DB 2; Length 432;
Best Local Similarity 39.0%; Pred. No. 1.1e-43;
Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;
Qy 33 ETTKEKSDNONQITDVSKISGLVNERKSEIMAAKADAKHFGMLMAIYTAGTVNDN 92

Db 3 ETDKGGKII-----RIFD-----NSFVKDRQAEIERA-----KNPDEFNTVLLTAGGTVDK 48
Qy 93 SFNOSWEAI-----QOLGALTG-GEITSVDSSTAIEGKYSILANTKNVWVLSGFQHD 147
Db 49 SFNOSIWEAVLEHVDQIEKTTNLDKRVSOETNNOSBLIKYKFLNGKNWVILTFQOQ 108
Qy 148 AFTREMLKIPEN-----KOLFTEKNIILIGIDW---TDTENVIPTRGYINLTYTEEAGWL 199
Db 109 EFPFELKQTDNSNGKKYSDLLAEKKVILIVAVDMDLSKEDKDLIKAGHFISSLYKTEEAGFI 168
Qy 200 AGYANASFLAKFPSPDTPKRSALVIGGISPAVVDFTLIGYLAGIKAMWLKNSDKTKTTT 259
Db 169 AGYASSKFLAYKFPDDEAKRTIAPFGGHHGAGVTDFTLIGYLAGIKAMWVNDPTGCVTISD 228
Qy 260 DKIEINLGFVDQDSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII--ANONDRYLIGV 317
Db 229 NNINIDTGF-ISNDKTATFINGIYNKS--SLVLPVAGSLTSSVDALIKSNKDTKYLIGV 285
Qy 318 DTDQSLVYTKTKNKFSTILKNLGYSVFSLDLYTKKSNRLAGFEFGK---KSATV 373
Db 286 DTDQSKIFSPA-TVFFTSIEKHLGRTIYQVLTIDMLKKEDSKFLGSFSLTNPANATV 344
Qy 374 YLGKDRFVDIADTSLSEGNKKLATEAISEAKKEFEKTKTIIPAEVAKTLEIPEM--P 430
Db 345 YKGISDPEVGVNSVTADADKVKAOEFLNEATADFKKOIQANPT-NYKSVLGIPMTLIND 403
Qy 431 DKOPDKOESL 441
Db 404 NDAKONEKASL 414

RESULT 8

AAP93343
ID AAP93343 standard; protein; 448 AA.

XX AAP93343;

XX 25-MAR-2003 (revised)

XX 01-AUG-1990 (first entry)

XX Gene encoding the 43 kd regression-associated antigen (RA Ag) of

XX Mycoplasma hyorhinis.

XX Tumour regression-associated antigens (RA Ag); in vivo imaging;

XX therapy monitoring; cancer therapy; Mycoplasma hyorhinis.

XX Mycoplasma hyorhinis.

XX Key Location/Qualifiers

XX Peptide 1..25

XX Protein 26..448

XX EPJ08265-A.

XX 22-MAR-1989.

XX 16-SEP-1988; 88EP-00308625.

XX 16-SEP-1987; 87US-00097910.

XX 11-DEC-1987; 87US-00131815.

XX (ITGE-) INT GENETIC ENG INC.

XX (INGE-) INGENE INT GENETIC.

XX (INGE-) INGENE INT GENETIC.

XX Fared GC, Sen A, Ghoshdasti P, Li A, Lee JH;

XX WPI; 1989-087638/12.

XX N-PSDB; AAN90684.

XX Tumour regression-associated antigens and antibodies - used in diagnostic

XX tests, monitoring course of therapy and for therapy in cancer patients.

XX Diaclosure; Page 7; 56pp; English.
PS Regression-associated antigens may be used in diagnostic tests, eg in
XX vivo imaging, for monitoring the course of therapy or for therapeutic
CC purposes, eg active immunisation protocols in cancer patients or drug
CC delivery systems by binding the drug to monoclonal or monospecific
CC polyclonal Ab showing specific immunoreactivity with the Ag. (Updated on
CC 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA
XX field.)
XX Sequence 448 AA;
SQ

Query Match 28.5%; Score 655; DB 1; Length 448;

Best Local Similarity 37.4%; Pred. No. 1e-40;

Matches 171; Conservative 89; Mismatches 149; Indels 48; Gaps 20;

14 GLVFPPLAIAT-ISAGCWDKETTKEESADNOKQITDVKISQUNVERKSEIIMAAADA 72

12 GTISTVASVATFVSCG-----ETDKGGKII-----RIFD-----NSFVKDRQAEIERA--- 54

73 NKHFGLMAIYTAGTVDNSFNOSWEAI-----QOLGALTG-GEITSVDSSTAIEGKY 127

55 -KNPDEFNTVLLTAGGTVDKXSFNOSIWEAVLEHVDQIEKTTNLDKRVSOETNNOSBLIKY 113

128 SSLANTKNVWVLSGFQHDGDAFTMLKIPEN-----KOLFTEKNIILIGIDW---TDTE 179

114 KNFLNGKNWVILTFQOQGEFPKFLKQTDNSNGKKYSDLLAEKKVILIVAVDMDLSKEDK 173

180 VIPTRGYINLTYTEEAGWLAVAGPLTEIFSDII--ANONDRYLIGV 239

174 LKAGHFISLKYTEEAGFIAGYASSKFLAYKFPDDEAKRTIAPFGGHHGAGVTDFTLAGF 233

240 LAGIKAMWLKNSDKTKITTDKIEINLGFVDQDSTKERLEQIASKDKPSTLLAVAGPLT 299

234 LAGIAKTNDNPTAKVTISDNINIDTGF-ISNDKTATFINGIYNKS--SLVLPVAGSLT 290

300 EIFSDII--ANONDRYLIGVTDQSLVYTKTKNKFSTILKNLGYSVFSLDLYTKKSN 357

291 SSVVDALIKSNKDKTKYLIGVTDQSKIFSPA-TVFFT-IEKHLGRTIYEVLTIDMLKKED 348

358 SRNLAGFEFGK---KSATYILGKDRFVDIADTSLSEGNKKLATEAISEAKKEFEKTK 413

349 S-FLGSFRSPFLTNAANAATYKGISD-FLGSNSVTADADKVKAOEFLN-ATADFKKOIQ 405

414 TTPAEVAKTLEIPEM-PDKOPDKOESLDKLINDIN 449

406 ANPT-NYKSVLGIPMLINDDAKNEKALNELIKIN 441

Db

RESULT 9

AA05332

ID AA05332 standard; peptide; 428 AA.

XX AA05332;

XX 25-JUN-1999 (first entry)

XX Inflammatory cytokine inducer.

XX Inflammatory cytokine inducer; thrombocytopenia; therapy.

XX Synthetic.

XX JPI1089582-A.

XX 06-APR-1999.

XX 19-SEP-1997; 97JP-00273649.

XX 19-SEP-1997; 97JP-00273649.

XX (CHUS) CHUGAI PHARM CO LTD.

XX

XX

XX

XX WPI; 1999-281057/24.
 DR N-PSDB; AAX33847.
 XX New inflammatory cytokine inducer gene and polypeptide - useful for
 PT treatment of thrombocytopenia.
 XX
 XX Claim 3; Page 16-17; 22pp; Japanese.
 XX
 CC This sequence represents the inflammatory cytokine of the invention. The
 CC inflammatory cytokine can be used in a drug, which is useful for the
 CC treatment of thrombocytopenia
 CC
 CC
 SQ Sequence 428 AA;
 Query Match 17.2%; Score 394.5; DB 2; Length 428;
 Best Local Similarity 26.8%; Pred. No. 4.2e-21;
 Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;
 QY 5 IKNNKFLGLVPEPLSAI-ATISAGCWDEKTEKESADNONKOITDVSIGSLVNERKS 63
 DB 1 MKSKSKILLGL-SPIALILPAVAVSCGNDES-----NISPEKDISKTTTNANGKQ 52
 QY 64 EIMAAKADANKHFGALNMAIYTAGTVNDNSFNOSSWEAIOQLATGCEITSVDSTAEI 123
 DB 53 VVKNAL- -LKLKPVLLTDEGKIDKSFNOSAFALAIKQIGIEINNVPS-SNF 106
 QY 124 EGKYSILANTNNKVVVLSGFQHGDAFTRMWKIPENKOLFTFKNIILIGIDMTDTENVIPT 183
 DB 107 ESAVNSALSGHKIWIWNGFRKHOOSIKQY--IDAHRELERNOIKIIGIDF-DIETRYK 163
 QY 184 GRVINTLYTKEAGWLAGVANASFLAKKPPSDPTKRSATVYIGGISPATVDFIAGYLAGI 243
 DB 164 --FYSIQFNKESAFITGVALASWLSQ--DESKRVASFGGAFGVITTFNMGFAKGI 218
 QY 244 KAMWLKNSDKTKI-TTDKIEINLGPVODSTKERLEQIAS-----KDKPESTLL 292
 DB 219 LYYNQKH--KSKIIYHSPVKLDSGF-----TAGKKNTVINNVLSSTPADVKKNPVIL 271
 QY 293 AVAGPLTEISDIIANONDRYLIGVDTQSLVYTKTKNKFETSLKWLGSVSVLSDL- 351
 DB 272 SVAGPAT--FETRYLANKGVYIGVSDQGM1--QDDRLITLVLKHIKQAVETLLDLI 327
 QY 352 -----YTKSNSNLAGFEFGKKSATVYIGIKDRFVDIADTSLGNDKUL-ATEAI 401
 DB 328 LEKEEGYKPVYVVDKADKTKSHFGTO-----KMKWGVABNHSNTEBOKXINKKI 379
 QY 402 SEAKKEFEKTKTIPAEVRKTLIEIPMP-KOPDKOESLDKLITDINK 450
 DB 380 KEAIIKMFKE---LPEDFVKYINSDKALKDGNKIDNVSERLAITSAINK 425
 RESULT 10
 ABP71709
 ID ABP71709 standard; protein; 428 AA.
 XX
 XX ABP71709;
 XX
 XX 04-APR-2003 (first entry)
 XX
 DE M161 antigen amino acid sequence.
 XX
 KM M161; antigen: cytotoxic T-lymphocyte inducer; CTL; cytostatic; cancer;
 KM anticancer agent.
 OS
 OS Mycoplasma fermentans.
 XX
 XX Key Location/Qualifiers
 FH M161 difference 121 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 FT M161 difference 163 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 FT M161 difference 163 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"

FT sequence"
 FT M161 difference 185 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 FT M161 difference 348 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 FT M161 difference 358 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 FT M161 difference 358 /note= "Encoded by TGA; unexplained stop codon in coding
 FT sequence"
 PN WO2002102402-A1.
 PD 27-DEC-2002.
 PF 13-JUN-2002; 2002WO-JP005916.
 PR 15-JUN-2001; 2001JP-00182250.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA Seya T, Matsumoto M, Naito K;
 PI WPI; 2003-167451/16.
 DR N-PSDB; ABZ59715.
 XX
 XX M161 antigen, its peptide fragment or their salt with e.g. superior CTL-
 PT inductive effect, applicable in anticancer agents combined with cancer
 PT antigen for preventing or treating cancer.
 XX
 XX Disclosure; Fig 1; 94pp; Japanese.
 XX
 CC The invention relates to cytotoxic T-lymphocyte (CTL)-inducers containing
 CC the M161 antigen, its partial peptide or their salt. The activity of CTL-
 CC inducers of the invention may be described as cytostatic. The antigen or
 CC its peptide fragment are applicable in anticancer agents for the
 CC prevention or treatment of cancer. The antigen and its derivative have
 CC superior CTL-inductive effect, giving anticancer agents that have low
 CC toxicity. The current sequence represents the M161 antigen amino acid
 CC sequence
 XX
 SQ Sequence 428 AA;
 Query Match 17.2%; Score 394.5; DB 6; Length 428;
 Best Local Similarity 26.8%; Pred. No. 4.2e-21;
 Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;
 QY 5 IKNNKFLGLVPEPLSAI-ATISAGCWDEKTEKESADNONKOITDVSIGSLVNERKS 63
 DB 1 MKSKSKILLGL-SPIALILPAVAVSCGNDES-----NISPEKDISKTTTNANGKQ 52
 QY 64 EIMAAKADANKHFGALNMAIYTAGTVNDNSFNOSSWEAIOQLATGCEITSVDSTAEI 123
 DB 53 VVKNAL- -LKLKPVLLTDEGKIDKSFNOSAFALAIKQIGIEINNVPS-SNF 106
 QY 124 EGKYSILANTNNKVVVLSGFQHGDAFTRMWKIPENKOLFTFKNIILIGIDMTDTENVIPT 183
 DB 107 ESAVNSALSGHKIWIWNGFRKHOOSIKQY--IDAHRELERNOIKIIGIDF-DIETRYK 163
 QY 184 GRVINTLYTKEAGWLAGVANASFLAKKPPSDPTKRSATVYIGGISPATVDFIAGYLAGI 243
 DB 164 --FYSIQFNKESAFITGVALASWLSQ--DESKRVASFGGAFGVITTFNMGFAKGI 218
 QY 244 KAMWLKNSDKTKI-TTDKIEINLGPVODSTKERLEQIAS-----KDKPESTLL 292
 DB 219 LYYNQKH--KSKIIYHSPVKLDSGF-----TAGKKNTVINNVLSSTPADVKKNPVIL 271
 QY 293 AVAGPLTEISDIIANONDRYLIGVDTQSLVYTKTKNKFETSLKWLGSVSVLSDL- 351
 DB 272 SVAGPAT--FETRYLANKGVYIGVSDQGM1--QDDRLITLVLKHIKQAVETLLDLI 327
 QY 352 -----YTKSNSNLAGFEFGKKSATVYIGIKDRFVDIADTSLGNDKUL-ATEAI 401

Dd	328	LEKEGKYPVVKDKKKADKXSHFGTQ-----XKKWIGVAENHFSSTEEQAQAKNNKI	379
Oy	402	SEAKKEFEKTKTIPAEVEYRKTLEIPEMPD-KDPKQOESLKLITDINK	450
Dd	380	KEAIKMFKE---LPDEFVKYINSDKALKDGNKINDVSRERLAISAINK	425
 RESULT 11 ADK68531			
ID	ADK68531	standard; protein; 428 AA.	
XX	ADK68531;		
AC			
XX			
XX	06-MAY-2004	(first entry)	
DT			
XX			
DE	Mycoplasma fermentans IL-X precursor protein #1.		
XX			
KM	Immunoregulatory factor; IL-X; lymphocyte proliferation.		
XX			
OS	Mycoplasma fermentans.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..24	
FT		/label=Signal_peptide	
FT	Protein	25..428	
FT		/note="Mature IL-X protein"	
XX			
PN	US6506892-B1.		
XX			
PD	14-JAN-2003.		
XX			
PF	29-OCT-1998;	98US-00182625.	
XX			
PR	29-OCT-1997;	97US-0063701P.	
XX			
PA	(WEBB/) WEBB A C.		
BA	(BLAZAR) BLAZAR B A.		
XX			
P1	Webb AC, Blazar BA;		
DR	WPI; 2003-401108/38.		
DR	N-PsDB; ADK68530.		
XX			
PT	New isolated polynucleotide, useful for augmenting proliferation of B and T lymphocytes.		
PS	Claim 25; SEQ ID NO 9; 42pp; English.		
XX			
CC	The invention provides novel immunoregulatory factor designated IL-X which has been isolated from Mycoplasma and polynucleotides encoding such polypeptides. The polynucleotide is useful for augmenting proliferation of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-X precursor protein.		
SQ	Sequence 428 AA;		
 Query Match 17.1%; Score 393.5; DB 7; Length 428; Best Local Similarity 26.8%; Pred. No. 5e-21; Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20			
Oy	5	IKANKFLGLGVFPLSAI-ATTSAGCWDKETTEKSKADNONKOTDVSKISGLVNERKS	63
Dd	1	MKSKSKIIILGL-SPIAAILPAVAASCGNNDSE-----NISFKEDISKYTTNANGKQ	52
Oy	64	EIMAAKADANKHFGLNMALIVTAGTVNONSFMQSSWEATOQLGTGEITSVDSTAEL	123
Dd	53	VVKONEL-----LKLPILTDEGKIDDKSFNQSAEALKAINKQTGISINNVS-SNF	106
Oy	124	BGYSSLANTNGVWLSGFOHGDAFTRLKIPENKQLFTEKNIIILIGIDMTDENVIP	183
Dd	107	ESAVYSALSAGHKIVNLNGFRQGSIKQY--IDAHRRELERNOIKITIGID-DIEITYEK	163
Oy	184	GRIYNLTYTEBAAGMLAGYANAFLAKKPPSPDKRSIAIVIGGSIAPAVTDIFAGYLAGI	243

Db	164	--FSLQGNIESAFTTGVALASWISLSEQ----	DESGRVVASFEQGAAPGVTTTNEGPAKFI	218
Qy	244	KAMLIKNSDKKTKTI-TTDKIEINLGFVDVDTSTKERLEQIAS-----	KDKPSTLL	292
Db	219	LYNQKH--KSKKIYHTSPVKLDSGF-----	TAGEKMTVINNVLSTFPADVKYNPHVIL	271
Qy	293	AVAGPLTIFSDIILNQMDRYLIGVDTPOSLYVTTKTKKFFSILIKNLGYSFVSLSDJ-		351
Db	272	SVAGPAT--FETVRLANKQGYVIGVDSQGM1--	ODKRIILTSVLKHIKQAVYEYLLDLI	327
Qy	352	-----YTKKSNSRMILAGFEFGKSKATVYLGIKDRFVIDIADISLSEGNDKL-ATEAI		401
Db	328	LEKEGKYFPYVVKDKKADKKKSHFETQ-----	KEKVIQVAENHFSTTEQAKIINNKI	379
Qy	402	SEAKKEPFEKTKTITPAEEVRKTLLEIPMPD-KQDPKQGESLDKLITDINK		450
Db	360	KEAIIKMFPE----LPEDFVKYINSDKALKDKGNKIDNVSRLEAIIAISAIWK		425
RESULT 12				
ID	ADK68538	standard; protein; 428 AA.		
AC	ADK68538;			
XX	DT	06-MAY-2004 (first entry)		
XX	XX	Mycoplasma fermentans IL-X precursor protein #2.		
DE	XX	Immunoregulatory factor; IL-X; lymphocyte proliferation.		
KM	XX	Mycoplasma fermentans.		
OS	XX			
XX	XX			
PH	FT	Key	Location/Qualifiers	
FT	Peptide	1..24	/label=Signal_peptide	
FT	Protein	25..428	/note="Mature IL-X protein"	
XX	XX	US6506892-B1.		
PN	XX	14-JAN-2003.		
XX	XX			
XX	PF	29-OCT-1998;	98US-00182625.	
XX	XX	29-OCT-1997;	97US-0063701P.	
XX	PR	(WEBB/) WEBB A C.		
XX	PA	(BLAZ/) BLAZAR B A.		
XX	PI	Webb AC, Blazar BA;		
XX	DR	WPI; 2003-401108/38.		
XX	DR	N-PSDB; ADK68535.		
PT	FT	New isolated polynucleotide, useful for augmenting proliferation of B and T lymphocytes.		
XX	XX			
PS	XX	Claim 26; SEQ ID NO 16; 42pp; English.		
CC	XX	The invention provides novel immunoregulatory factor designated IL-X		
CC	XX	which has been isolated Mycoplasma and polynucleotides encoding such		
CC	XX	polypeptides. The polynucleotide is useful for augmenting proliferation		
CC	XX	of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-		
CC	XX	X precursor protein.		
SQ	XX	Sequence 428 AA;		
Query Match	17.0%	Score 390.5;	DB 7;	Length 428;
Beet Local Similarity	26.8%	Pred. NC. 8.4e-21;		
Matches 126;	Conservative 100;	Mismatches 175;	Indels 69;	Gaps 20

PI Webb AC, Blazar BA;
XX WPI; 2003-401108/38.
DR N-PSDB; ADK68530.
XX
XX New isolated polynucleotide, useful for augmenting proliferation of B and
PT T lymphocytes.
XX
XX Claim 12; SEQ ID NO 14; 42pp; English.
XX
XX The invention provides novel immunoregulatory factor designated IL-X
CC which has been isolated Mycoplasma and polynucleotides encoding such
CC polypeptides. The polynucleotide is useful for augmenting proliferation
CC of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-
CC X mature protein.
XX
SQ Sequence 404 AA;

Query Match 16.7%; Score 384.5; DB 7; Length 404;
Best Local Similarity 26.9%; Pred. No. 2.2e-20;
Matches 116; Conservative 89; Mismatches 155; Indels 71; Gaps 17;

QY 59 NERKEIMAAKADNKH-----FGLMAIVTAGTVNDNSFNQSWEMAI 102
DB 3 NNDSSNISFKKIDSKYTTTANANGKOVVNAELIKLKPILITDEGKIDDSFNQSAFEL 62

QY 103 QOLGALTGEITSVDSSTAELEGKYSLSANTKNVWVLSGFQHDGDAFTRLKIPENKOLF 162
DB 63 KAIKKQGIIEINNEPS-SNFESAYNSALSAGHKIWLNGFKHQOSIKQY--IDAYREEL 119

QY 163 TEKNIILIGIDWDTENVIPTRGRYINLTYTEAGMWLAGYANASFLAKKPPSDPTKSAI 222
DB 120 ERNOIKIKIGIDF-DIETEVKM--FYSLOFNKIKESAFITGYAIAWSLSEQ--DESKRVVA 173

QY 223 VIGGISPAYTDFIAGYLAGIKAMNLIKNSDKTKI--TTDKIEINLGFVDQDTSKERLEQ 281
DB 174 SFGGGAFFGVTTFEGGAKGILYNNQK--KSKKIYHTSPVKLDSGF-----TAGKKMNT 226

QY 282 IAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVDTQSLVYTKTKNK 331
DB 227 VINNVLSSTPADVKNPHVILSVAGPAT--FETVRLANKGGYVIGVSDQGM--QDKDR 282

QY 332 FFTSILKNLGYSVFSVLSL-----YTKKSNRNLAGFEGKKATVYLGIDRF 381
DB 283 ILTSLVLRHAKQAVETLIDLILEKEEGKPYVVDKKDKKWSHFQTQ-----KEKW 334

QY 382 VDIADTSLGNDKKL-ATEAISEAKKEFEKTKTIIPAEVRKTLIEIPMP-KQPDQOE 439
DB 335 IGVAENHFSTNEBQAKINNKIKKAIKMFKE---LPEDFVKYINSKALKDKGNKIDNVSE 390

QY 440 SLDKLITDINK 450
DB 391 RLEAIIISAINK 401

RESULT 15
ADK68537
ID ADK68537 standard; proteini; 404 AA.
XX
XX ADK68537;
XX
XX 06-MAY-2004 (first entry)
XX
XX Mycoplasma fermentans IL-X mature protein #2.
XX
XX Immunoregulatory factor; IL-X; lymphocyte proliferation.
XX
XX Mycoplasma fermentans.
XX
XX US6506892-B1.
XX
XX 14-JAN-2003.
XX

PF 29-OCT-1998; 98US-00182625.
XX
XX 29-OCT-1997; 97US-0063701P.
XX
XX (WEBB/) WEBB A C.
PA (BLAZ/) BLAZAR B A.
XX
XX Webb AC, Blazar BA;
XX
XX WPI; 2003-401108/38.
DR N-PSDB; ADK68535.
XX
XX New isolated polynucleotide, useful for augmenting proliferation of B and
PT T lymphocytes.
XX
XX Claim 13; SEQ ID NO 15; 42pp; English.
XX
XX The invention provides novel immunoregulatory factor designated IL-X
CC which has been isolated Mycoplasma and polynucleotides encoding such
CC polypeptides. The polynucleotide is useful for augmenting proliferation
CC of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-
CC X mature protein.
XX
SQ Sequence 404 AA;

Query Match 16.6%; Score 381.5; DB 7; Length 404;
Best Local Similarity 26.9%; Pred. No. 3.7e-20;
Matches 116; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

QY 59 NERKEIMAAKADNKH-----FGLMAIVTAGTVNDNSFNQSWEMAI 102
DB 3 NNDSSNISFKKIDSKYTTTANANGKOVVNAELIKLKPILITDEGKIDDSFNQSAFEL 62

QY 103 QOLGALTGEITSVDSSTAELEGKYSLSANTKNVWVLSGFQHDGDAFTRLKIPENKOLF 162
DB 63 KAIKKQGIIEINNEPS-SNFESAYNSALSAGHKIWLNGFKHQOSIKQY--IDAYREEL 119

QY 163 TEKNIILIGIDWDTENVIPTRGRYINLTYTEAGMWLAGYANASFLAKKPPSDPTKSAI 222
DB 120 ERNOIKIKIGIDF-DIETEVKM--FYSLOFNKIKESAFITGYAIAWSLSEQ--DESKRVVA 173

QY 223 VIGGISPAYTDFIAGYLAGIKAMNLIKNSDKTKI--TTDKIEINLGFVDQDTSKERLEQ 281
DB 174 SFGGGAFFGVTTFEGGAKGILYNNQK--KSKKIYHTSPVKLDSGF-----TAGKKMNT 226

QY 282 IAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVDTQSLVYTKTKNK 331
DB 227 VINNVLSSTPADVKNPHVILSVAGPAT--FETVRLANKGGYVIGVSDQGM--QDKDR 282

QY 332 FFTSILKNLGYSVFSVLSL-----YTKKSNRNLAGFEGKKATVYLGIDRF 381
DB 283 ILTSLVLRHAKQAVETLIDLILEKEEGKPYVVDKKDKKWSHFQTQ-----KEKW 334

QY 382 VDIADTSLGNDKKL-ATEAISEAKKEFEKTKTIIPAEVRKTLIEIPMP-KQPDQOE 439
DB 335 IGVAENHFSTNEBQAKINNKIKKAIKMFKE---LPEDFVKYINSKALKDKGNKIDNVSE 390

QY 440 SLDKLITDINK 450
DB 391 RLEAIIISAINK 401

Search completed: December 18, 2004, 01:22:19
Job time: 118.524 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:16:03 ; Search time 30.9611 Seconds
(without alignments)
966.032 Million cell updates/sec

Title: US-09-676-249D-2

Perfect score: 2299
Sequence: 1 MKKKIKWKKFLGLVLPPLS.....KPPDKQESDILITDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.rep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.rep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.rep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.rep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUB_COMB.rep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.rep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393.5	17.1	428	US-09-182-625F-9	Sequence 9, Appl1
2	390.5	17.0	428	US-09-182-625F-16	Sequence 16, Appl1
3	384.5	16.7	404	US-09-182-625F-14	Sequence 14, Appl1
4	381.5	16.6	404	US-09-182-625F-15	Sequence 15, Appl1
5	225	9.8	363	US-09-583-110-2686	Sequence 2686, Ap
6	203	8.8	328	US-08-961-083-8	Sequence 8, Appl1
7	203	8.8	328	US-09-536-784-8	Sequence 8, Appl1
8	203	8.8	328	US-09-107-532A-5084	Sequence 5084, Ap
9	199.5	8.7	362	US-09-134-000C-6004	Sequence 6004, Ap
10	182	7.9	339	US-09-071-035-80	Sequence 80, Appl1
11	182	7.9	361	US-09-071-035-78	Sequence 78, Appl1
12	182	7.9	375	US-09-134-000C-6005	Sequence 6005, Ap
13	179	7.8	341	US-08-396-957A-5	Sequence 5, Appl1
14	173	7.5	166	US-09-182-625F-6	Sequence 6, Appl1
15	162	7.0	339	US-08-396-957A-4	Sequence 4, Appl1
16	149	6.5	867	US-09-540-232-2676	Sequence 2676, Ap
17	144	6.3	353	US-08-313-412-1	Sequence 1, Appl1
18	141.5	6.2	889	US-09-336-447A-15	Sequence 15, Appl1
19	141.5	6.2	889	US-09-952-267B-15	Sequence 15, Appl1
20	132	5.7	553	US-09-134-001C-2974	Sequence 2974, Ap
21	131.5	5.7	518	US-08-973-462-8	Sequence 8, Appl1
22	130.5	5.7	518	US-09-248-796A-18238	Sequence 18238, A
23	126	5.5	557	US-09-134-000C-5061	Sequence 5061, Ap
24	122.5	5.3	2285	US-09-308-375-2	Sequence 2, Appl1
25	122	5.3	2411	US-09-268-347-36	Sequence 36, Appl1
26	121.5	5.3	834	US-09-187-999-11	Sequence 11, Appl1
27	121.5	5.3	1104	US-08-923-992A-4	Sequence 4, Appl1

28	120	5.2	878	4	US-09-540-236-3401	Sequence 3401, Ap
29	119.5	5.2	1861	2	US-08-790-912-4	Sequence 4, Appl1
30	119	5.2	892	2	US-09-336-447A-5	Sequence 5, Appl1
31	119	5.2	892	4	US-09-952-267B-5	Sequence 5, Appl1
32	117.5	5.1	1073	3	US-09-074-658-75	Sequence 75, Appl1
33	117.5	5.1	1073	4	US-09-206-942-49	Sequence 49, Appl1
34	117.5	5.1	1079	4	US-09-206-942-47	Sequence 47, Appl1
35	117	5.1	1912	1	US-08-409-995-4	Sequence 4, Appl1
36	117	5.1	1912	3	US-08-685-467-4	Sequence 4, Appl1
37	117	5.1	2353	3	US-09-377-155-33	Sequence 33, Appl1
38	117	5.1	2353	3	US-08-913-942-4	Sequence 4, Appl1
39	117	5.1	2353	3	US-09-669-974-33	Sequence 33, Appl1
40	117	5.1	2353	4	US-09-797-862-33	Sequence 33, Appl1
41	117	5.1	2354	4	US-09-684-707-4	Sequence 4, Appl1
42	117	5.1	2354	3	US-09-268-347-47	Sequence 47, Appl1
43	116.5	5.1	873	3	US-09-336-447A-13	Sequence 13, Appl1
44	116.5	5.1	873	4	US-09-952-267B-13	Sequence 13, Appl1
45	115.5	5.0	1164	3	US-08-923-992A-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-182-625F-9
; Sequence 9, Application US/09182625F
; Patent No. 6506892
; GENERAL INFORMATION:
; APPLICANT: Webb, Andrew C.
; TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gr
; FILE REFERENCE: BLAZ-101XCI
; CURRENT APPLICATION NUMBER: US/09/182, 625F
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/063, 701
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 9
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Mycoplasma fermentans
US-09-182-625F-9

Query Match      17.1%   Score 393.5; DB 4; Length 428;
Best Local Similarity 26.8%; Pred. No. 1.7e-27;
Matches 126; Conservative 101; Mismatches 174; Indels 69; Gaps 20;

QY      5 IKMKKFLGLVLPPLSAI-ATISAGCWDKETTKEKSDNQNKQITDVSKISGLVNERKS 63
      1 MKSKSKITLGL-SPIALIPAVAVSCGNDES-----NISFKEDISKVTTNANGKO 52
      64 EIMAAKDANKHGLMAIYTAGTVNDSFNQSSWAIQQLGALTGCEITSVDSIAEL 123
      53 VVKNAEL-----LTKLPILITDEGKIDKSFNSAFALKAINKQIGIENNVPS-SNF 106
QY      124 EGKSSLANTNNKVVVSGFOHGAFTRLMKIPENKQLTFEKNIIIGIMTDENIYPT 183
      107 ESNALSLSLGHKKIKWVNGFRHQOSIKQY-IDAHRELEBNQIKIGIDF-DLETXYK 163
      184 GRYINLTYTEAGWGLAVANASFLAKKFPSPDKSAIYIGGISPAVTDIAGVLAGI 243
      164 --FVSLQFNKESAFITGVAIASGLSQ---DESKRVAASFQGGAPGVTTFNGFAFKI 218
      244 KAMTLKSDKTKI-TTDKLEINIGFVQVOTSTRELEQIAS-----KDKESTLL 292
      219 LYYNQKH--KSKSKYHTSPVKLDSGF-----TAGERKNTVINNVLSSTPADVKYNPVIL 271
QY      293 AVAGPLLEIRSDIANQNDRIYLGVDPTQSLVYTKTKNGFFTSILKVLGYSVFVLDL 351
      272 SVAGPAT--ETVYRLANKGVVIGVSDQGI--QDKDRILTSVLKIKKQAVETFLDLI 327
      352 -----YTKKSNSHNLGAFEPGKSKATVYIGIKDRFVDIADTSLGNDKGL-ATEAI 401
```

```
Db      328 LEKEEGKPPYVVKKKADKKMSHFQTQ-----KEKMWIGVANNHNSNTSEOKAKINKKI 379
Qy      402 SEAKKEFEKTKTTPAEVRKLTLEIPEMPD-KQPDKQOESLDTITDINK 450
      380 KEALIKMKE-----LPEDFVKYINSDKALKDGNKIDNVSERLEAITSAINK 425

RESULT 2
US-09-182-625F-16
; Sequence 16, Application US/09182625F
; Patent No. 6506892
; GENERAL INFORMATION:
; APPLICANT: Webb, Andrew C.
; APPLICANT: Blazar, Beverly A.
; TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gr
; FILE REFERENCE: BLAZ-101XC1
; CURRENT APPLICATION NUMBER: US/09/182,625F
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/063,701
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 16
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Mycoplasma fermentans
US-09-182-625F-16

Query Match      17.0%; Score 390.5; DB 4; Length 428;
Best Local Similarity 26.8%; Pred. No.3.3e-27;
Matches 126; Conservative 100; Mismatches 175; Indels 69; Gaps 20;

Qy      5 IKMKPLGLGLVPLPSAI-ATISAGCWDKETTKEKSDNONKOITDVSKISGLVNERKS 63
      1 MKSKKTLILGL-SPIAAILPAVAVSCGNDES-----NISFKKOISKYTTTNANGKO 52
Db      64 EIMAAKADANKHFGIMNAIYTAGTVNDNSFNQSSWEAIIQOLGALTGGEITSVDSTAEI 123
      53 VVKNAEL-----LKLKPLILITDEGKIDKSFNQSFAFEALKAINKQGTGIEINNVPS-SNF 106
Qy      124 EGKSSSLANTNKVWVLSGFQHGDAFTRWLKIPEKNOIFTEKNIILIGIDWTDENVIPT 183
      107 ESNVNSALSAGHKIWLNGFPGQOSIKOY-IDAYRELEENQIKITIGIDF-DIETEKW 163
Qy      184 GRVYNLYTKEEAGMLAGYANASFLAKKFPSPDKSAIIVGGGISPAVTDPIAGYLAGI 243
      164 --FYSLOFNIKESAFITGYAIAWSLSEQ--DESKRVVASFGGAFPGVTTFNNGFAKGI 218
Qy      244 KANWLKNSDKTKI-TTDKIEINLGFVDVQDTSTYERLEQIAS-----KDKPSTLL 292
      219 LLYNQKH-KSSKIYHTSPVKLDSGF-----TAGKKNMTVINNVLSSTPADVKYNPHVIL 271
Qy      293 AVAGLEIFSDILIANQNDRLIGVDTPDOSLYTKTKNKFFTSILKXUGSVFSLADL 351
      272 SVAGPAT--FETVLANKGQYIVGVDSQGM--QDKRIITFVLKHKIQAVYETLIDL 327
Qy      352 -----YTKSNSRNLAGFEFGKSAFVYLGIKDFVVDIADTSLGNDKKI-ATEAI 401
      328 LEKEEGKPPYVVKKKADKKMSHFQTQ-----KEKMWIGVANNHNSNTSEOKAKINKKI 379
Qy      402 SEAKKEFEKTKTTPAEVRKLTLEIPEMPD-KQPDKQOESLDTITDINK 450
      380 KEALIKMKE-----LPEDFVKYINSDKALKDGNKIDNVSERLEAITSAINK 425

RESULT 3
US-09-182-625F-14
; Sequence 14, Application US/09182625F
; Patent No. 6506892
; GENERAL INFORMATION:
; APPLICANT: Webb, Andrew C.
; APPLICANT: Blazar, Beverly A.
```

```
; TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gr
; FILE REFERENCE: BLAZ-101XC1
; CURRENT APPLICATION NUMBER: US/09/182,625F
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/063,701
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 14
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Mycoplasma fermentans
US-09-182-625F-14

Query Match      16.7%; Score 384.5; DB 4; Length 404;
Best Local Similarity 26.9%; Pred. No.1.1e-26;
Matches 116; Conservative 89; Mismatches 155; Indels 71; Gaps 17;

Qy      59 NERKSEIMAAKADANKI-----FGLMNAIYTAGTVNDNSFNQSSWEAI 102
      3 NDNESNISFKKIDSKYTTTNANGKQVKNNAELKLKPLITDEGKIDKSFNQSFAFEAL 62
Qy      103 QOLGALTGGEITSVDSTAELEGKSSSLANTNKVWVLSGFQHGDAFTRWLKIPEKNOIF 162
      63 KAIKQGTGIEINNVPS-SNFEKSNVNSALSAGHKIWLNGFPGQOSIKOY-IDAHREEL 119
Qy      163 TEKNIILIGIDWTDENVIPTGRVYNLYTKEEAGMLAGYANASFLAKKFPSPDKSAI 222
      120 ERNQIKITIGIDF-DIETEKW--FYSLOFNIKESAFITGYAIAWSLSEQ--DESKRVVA 173
Qy      223 VIGGISPATYTDPIAGIAGIKANLKNSEKTKI-TTDKIEINLGFVDVQDTSTYERLEQ 281
      174 SPFGGAPPGVYTFNNGFAKGIILYNNQKH-KSSKIYHTSPVKLDSGF-----TAGKKNMT 226
Qy      282 IAS-----KDKPSTLLAVAGPTEIFSDILIANQNDRLIGVDTPDOSLYTKTKNK 331
      227 VINNVLSSTPADVKYNPHVILSVAGPAT--FETVLANKGQYIVGVDSQGM--QDKR 282
Qy      332 FPTSILKXUGSVFSLSDI-----YTKSNSRNLAGFEFGKSAFVYLGIDRF 381
      283 ILTSVLKHKIQAVYETLIDLILEKEEGKPPYVVKKKADKKMSHFQTQ-----KEK 334
Qy      382 VDIADTSLGNDKKI-ATEAISEKKEFEKTKTTPAEVRKLTLEIPEMPD-KQPDKQOE 439
      335 IGAENHFSNTEBQAKINKKIKEALIKMKE-----LPEDFVKYINSDKALKDGNKIDNVSE 390
Qy      440 SLDKLITDINK 450
      391 RLEAITSAINK 401

Db

RESULT 4
US-09-182-625F-15
; Sequence 15, Application US/09182625F
; Patent No. 6506892
; GENERAL INFORMATION:
; APPLICANT: Webb, Andrew C.
; APPLICANT: Blazar, Beverly A.
; TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gr
; FILE REFERENCE: BLAZ-101XC1
; CURRENT APPLICATION NUMBER: US/09/182,625F
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: US 60/063,701
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 15
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Mycoplasma fermentans
US-09-182-625F-15

Query Match      16.6%; Score 381.5; DB 4; Length 404;
```


Best Local Similarity 26.9%; Pred. No. 2e-26;
Matches 116; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

QY NERSEIIMAKADANKI-----FGLMAIYTAGTNDNSFNSSWEAI 102
DB 3 NDESNISFKEKDISKTYTTNANGKQVNAELLKPIILTDGKIDDSFNSSAREAL 62
QY 103 QOQALTCGETTSVDSSTAELEGYSILANTNNKVVVLSGFQHDATFRMLKIPENKOLF 162
DB 63 KAINKQCIIEINNVPS-SNFESAYNSALSGHKIWLNGFKHOOSIKOY--IDAYEEL 119
QY 163 TEKNIILIGIDWTDENVIPTRGRYINLTYKTEAGMLAGVANASFLAKPSPDPTKSAI 222
DB 120 ERNOIKIIGIDF-DIETEVKW--FYSLOFNIKESAFITGVAIASWLSQ--DESKRVVA 173
QY 223 VIGGIGIPATYDFLAGLACIKAMNLSKDKTKI--TTDKIENLGPVODTSTKERLEQ 281
DB 174 SFGGAPFPVYTFEFGAKGLIYNOQH--KSKIIYHSPVKLDSGF----TAGEKMT 226
QY 282 IAS-----KDPSTLLAVAGPLTEIFPSDIIANONDRYLIGVDTQSLVYTKTKNK 331
DB 227 VINNVLSSTADVKNHVLISVAGPAT--FETVRLANKGOYVIGVSDQGM--QDKDR 282
QY 332 FTSILKNLGSYVPSVLSL-----YTKKSNRNLAGFEFGKKSATVYLGIKDRF 381
DB 283 ILTSVLKHIQAVYETLLDILEKEGKYPVVDKADKKWSHFGTQ-----KEXW 334
QY 382 VDIADTLEGNDKL-ATEAISEAKKEFEETKTIIPAEFRKTEIEMPD-KOPDQOE 439
DB 335 IGAENHFSNTEBQAKINNKIKAIKMFKE---LPEDFYKINSKALKGKNKIDNVSE 390
QY 440 SLDKLITDINK 450
DB 391 RLENIISAINK 401

RESULT 5

US-09-583-110-2686
Sequence 2686, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PAT400-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 2686
LENGTH: 363
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-2686

Query Match 9.8%; Score 225; DB 4; Length 363;
Best Local Similarity 24.8%; Pred. No. 3.9e-12;
Matches 110; Conservative 61; Mismatches 150; Indels 122; Gaps 21;

QY 1 MKKKIKNNKGLGLVPLSAIATIS-AGCWDKETTKEEKADNONKOITDVSKISGLVN 59
DB 14 MNKK-----QWLGGLV---AVALVGLAACNRRSSRNAASSSDVYTK-----52
QY 60 ERKEIIMAAKADANKHGLNMAIYTAGTVNDNSFNSSWEAIQOLG---LTGETITSV 116
DB 53 -----NAIVTDGVDKDSFNSSWAMEGLQAWGKEHNLKONGFTY 92
QY 117 DSSSTAELE-----GKYSILANTNNKVVVLSGFQHDATFRMLKIPENKOLFTEKNI 167

DB 93 FOSTSEADYANNLQQAAGSYNLI-----FGVGFALHNAVEBAKHTDLNY 138
QY 168 IILIGIDWTDENVIPTRGRYINLTYKTEAGMLAGVANASFLAKPSPDPTKSAIIVGCG 227
DB 139 VLLIDVVKDKNV-----ASVFPADNBSGLAGVAAK-----TTTKOVGVGG 183

QY 228 I-SPAVDPLFAGYLAGIKAMNLSKDKTKITTDKIEINLGFVODTSTKERLEOIASKD 286
DB 184 IESVIRFEPAGFRAGV-----ASVDSIRVQVDYAG-SFG-DAAKGTIAAQAAGAD 236
QY 287 KPSTLLAVAGPL-TEIFPSDIIANONDR-----YLIGVDTQSL-VYT---KTQKFF 333
DB 237 ---IVYQVAGTGAGVAFKAASLNSHRENEKVVIGVDRDOEAKGKYSKDGKESNFVL 293
QY 334 TSILKNLGSYVPSVLSLTYTKKSNRLAGFEFGKKSATVYLGIKDRFVDIADTSLGND 393
DB 294 VSTLKQVGTIVKDI-----SNKAEKGFPGSQ--VIVYSKDKGLAVLNTLSSEG 342
QY 394 KKLATEAISEAKKEFEETKTIIP 416
DB 343 KK-----AVEDAKAKILDGSVKVP 361

RESULT 6

US-08-961-083-8
Sequence 8, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-8

Query Match 8.8%; Score 203; DB 3; Length 328;
Best Local Similarity 25.7%; Pred. No. 3.4e-10;
Matches 98; Conservative 56; Mismatches 144; Indels 84; Gaps 19;

QY 61 RKSEIIMAKADANKHGLNMAIYTAGTVNDNSFNSSWEAIQOLG---LTGETITSVD 117

Db 3 RSSRNAASSDVK---TKAAIVTDGTGVDKSFNQSAMEGLQAMGKEHNLKONGFTYF 58
Qy 118 SSTAEL-----GKYSILANTKNKVVLSGFQGDAPTRMLKIPENKOLFTEKNII 168
Db 59 QSTSEADVANNLQQAAGSYNLI FGV-----GALNNA-----VKDAKETHDNLNV 104
Qy 169 ILIGIDWTDENVITPTGRYINLTYTEAGMLAGYANASFLAKKPPSDPTKSAIVIGGI 228
Db 105 LIDVIVDQKXV-----ASVTFADNESGYLAGYAAAK-----TTTKQVGFVGGI 149
Qy 229 -SPAVIDPIAGYLAGIKAMNLKNSDKKTKITTDKIEINLGFVDPTSTKERLEQIASDK 287
Db 150 ESEVISFEAGFKAGV-----ASVDPSIKQVDYAG-SFG-DAKAKTIAAQAAGAD- 201
Qy 288 PSTLLAAGPL-TEIFSDIANQDR-----YLIGVDTQSL--VYT---KTKNKFT 334
Db 202 --IYQVAGGTGAGVFAEAKSLNESRPNENKWWI GVDRODEAGKYTSKDGKSNFVLV 259
Qy 335 SILKNLGYSVFVSLDYTKKNSRNLAGEFEGKKSATVYIGIKDRFVDIADTSLSENDK 394
Db 260 STLKQVGTIVKDI-----SNKAERGEFPGGQ--VIYVSLKDKGVDLAVTNLSEBK 308
Qy 395 KLATEAISEAKKEFEKTKTIP 416
Db 309 K---AVEDAKAKILDSGVKVP 326

RESULT 7

US-09-536-784-8
; Sequence 8, Application US/09536784
; Patent No. 6573082

GENERAL INFORMATION:

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: P8340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-536-784-8

Query Match 8.8%; Score 203; DB 4; Length 328;
Best Local Similarity 25.7%; Pred. No. 3.4e-10;

Matches 96; Conservative 56; Mismatches 144; Indels 84; Gaps 19;
Qy 61 RKSEIMAKADANKHFGIANNAIYTAGTVNDNSFNQSWEAIQOLA--LTGGITSV 117
Db 3 RSSRNAASSDVK---TKAAIVTDGTGVDKSFNQSAMEGLQAMGKEHNLKONGFTYF 58
Qy 118 SSTAEL-----GKYSILANTKNKVVLSGFQGDAPTRMLKIPENKOLFTEKNII 168
Db 59 QSTSEADVANNLQQAAGSYNLI FGV-----GALNNA-----VKDAKETHDNLNV 104
Qy 169 ILIGIDWTDENVITPTGRYINLTYTEAGMLAGYANASFLAKKPPSDPTKSAIVIGGI 228
Db 105 LIDVIVDQKXV-----ASVTFADNESGYLAGYAAAK-----TTTKQVGFVGGI 149
Qy 229 -SPAVIDPIAGYLAGIKAMNLKNSDKKTKITTDKIEINLGFVDPTSTKERLEQIASDK 287
Db 150 ESEVISFEAGFKAGV-----ASVDPSIKQVDYAG-SFG-DAKAKTIAAQAAGAD- 201
Qy 288 PSTLLAAGPL-TEIFSDIANQDR-----YLIGVDTQSL--VYT---KTKNKFT 334
Db 202 --IYQVAGGTGAGVFAEAKSLNESRPNENKWWI GVDRODEAGKYTSKDGKSNFVLV 259
Qy 335 SILKNLGYSVFVSLDYTKKNSRNLAGEFEGKKSATVYIGIKDRFVDIADTSLSENDK 394
Db 260 STLKQVGTIVKDI-----SNKAERGEFPGGQ--VIYVSLKDKGVDLAVTNLSEBK 308
Qy 395 KLATEAISEAKKEFEKTKTIP 416
Db 309 K---AVEDAKAKILDSGVKVP 326

RESULT 8

US-09-107-532A-5084
; Sequence 5084, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Walcham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5084:

SEQUENCE CHARACTERISTICS:
LENGTH: 374 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein


```

QY      61 EKSEIEMAKADANKHFGILMIAVLTGAGTJNDSPNOSMEALIOQLA-----LTGCB----- 112
Db      7 KTBEGGGKGD-----AHSAVIITDTGGVDDSPNOSMEGIQAMCKEHDLPBGSGCYAV 62
QY      113 ITSVDSS--TAELEGRYSGLANTNRKNWVLSGFHGDATFRMLKIPENKOLFTEKNIIL 170
Db      63 IQSNDAADYTTNIDQAVSKFNTPFGI-----GYLKDA-----ISSAADANPDNFWVL 112
QY      171 GIDWDTENVLPTRKYNILTYKTEBAGMLAGYANMSFLAKKPEPSDPTKSAIYIGGJISB 230
Db      113 DDQIDGKKNV-----VSATFRDNEAAYTAGYAAA-----NETKTNKRGFVGGEBS 157
QY      231 AVTD-FIAGYLAGIKAMMLKNSDKTKITTDKIEINLGFVDQDTSRKERLEQJASDKDPS 289
Db      158 VVIDRFQGFGEFV-ADAAKELGKETVDT-----KTAASPADPR 196
QY      290 TLLAVAGFLTEIFSDII-----ANQNR-YLIGVDTQSL--VYT 326
Db      197 KGKMLAAMYQNGVDDIIFASGATGCGVQEAKDNLSESGSGDKWVIGVDRDDADAGKY- 255
QY      327 KTK-----NKEFTSLKNLGYSVFESVLDLYTKKSNSRNLAGFEFGKSGATVYLGIKDRF 381
Db      256 KTKGKEDNFETLSLKGAGTAVODIANALEDK-----FPGGHLVY-GKDDG 304
QY      382 VDIADPTEGNDKCLATEAISAKKEFEKTKTPIAEVFRKLTLEIPEMD 431
Db      305 VDLTDGYL-NDKT-----KAVVTAKDKVLSGDGVKVEKBE 339

```

```

RESULT 11
US-09-071-035-78
; Sequence 78, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-78

```

	Query Match	7.9%;	Score 182;	DB 4;	Length 361;	
	Best Local Similarity	25.4%;	Pred. No. 3.4e-08;			
	Matches	104;	Conservative	51;	Mismatches 139;	Indels 116; Gaps 20;
Qy	61	RKSSIMAAKADANKHFGIINMAIVTAGGTVDNDSFNQSWSEAIQQLGA--LTGE-----	112			
Db	29	KTAAESGGGKGGA-----ASHAVIITDITGVDGDDKSPNOSWBEIQLMWGKEHNDPEGSKGAY	84			
Qy	113	ITSDVSS--TAELEGKSSLANTKHKNWVLVSGFHGDAFTRWLAIPEKQLFTEKNIIIL	170			
Db	85	IQSDNDAADYTTNIDQAVSSKFNITFGI---GYLLKDA-----ISSAADANPDINFLVI	134			
Qy	171	GIDWTDENVIPTGRYINLTYTEAGWLAGYANASFLAKKFPSDPTKRSALVIGGGISP	230			
Db	135	DDQIDGKKNV-----VSATFRDNEAVALAGVAAA-----NETKTNKGVFGVBEG	179			
Qy	231	AVTD-FLAGIYLAGIKAMNLIKNSDDKTKITTDPKIEINLGFVDYDQISTYKERLEQIAKSKKPS	289			
Db	180	VIDRFQAGFEKGV-ADAAPKELGEKITVDT-----RYAASFADPA	218			
Qy	290	TLVAVAGLPTIEPSDII-----ANQNDR-VLIGVDPDQSL--VYT	326			
Db	219	KGKALLAAMYANGVDIIFHSGATGGGVPFGFAKOLNBSGSDKAVWVICVDDDDQADAGKY	277			
Qy	327	KTK-----NKFETSIKNLGYSPFVSLDLYTKKSNBSNLAGFEFGKKASATVYIGIDRF	381			
Db	278	KTKQKGEKDNFLTSTLKGVTGAVODIANRALEDK-----FPGEGHLVY-GLXDG	326			
Qy	382	VDIADTSLGNDKKLATAEALSEAKKEPEPEKTKTIPAEVRKTLRIPEKPD	431			
Db	327	VDLIDGCVL-NDKT-----KEAVKTKADKVIISDVAVPEKPE	361			

```

RESULT 12
US-09-134-000C-6005
; Sequence 6005, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIORITY FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIORITY FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6005
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6005

```

	7.9%; Score 182; DB 4; Length 375; Best Local Similarity 25.4%; Pred. No. 3.66-08; Matches 104; Conservative 51; Mismatches 139; Indels 116; Gaps 20;
QY	61 RKSEIMAKADANKHFGIINMAIVTAGGTVDNDSFNQSSWEAIIQQLGA---LTGEE----- 112
Db	43 KTAESGGGKGDA---AHSAAVILITDTGGVDGDKSFENQSSWEGLQMWGKEHNDIPESKGYAY 98
QY	113 ITSDSS--TALEEGKXSLANTKONWVLSFGHGAFTWMLKIPENKOLFTEKNITIL 170
Db	99 IQSDNADADYTTNTIQAVSSKFNITFGI---GYLLKDA-----ISSAADNPPTNFYLI 148
QY	171 GIDWTDTEENVIPTRGRYINLTYTEEAGWLAGYANASFLAKFPSPDPTKRSAAIVIGGISP 230
Db	149 DDQIDGKKNV-----VSATFRDNEAAYLAGVAAA-----NETTKNVKGVFGGEEG 193
QY	231 AVTD-PLAGYIAGIKANMLKNSDKKTKITTDKIEINLGFVDYDTSTKERLEQIAKSKPS 289
Db	194 VVIRFQAGEFGKV-ADAARELGEIETIVDT-----RYAASFADPA 232

QY 290 TLAVAGPLTEIFEDII-----ANQDR-YLIGVDTQSL--VYT 326
 Db 233 KGAALAAAMVONGVDIIFHAGSAGTGGVFGAOKDLNBSGSDKXVWVGVDQDQADGKX- 291
 QY 327 KTK-----NKFPSILKNLGYSPVLSDLTYTKXSNBNLAGFPFGKSAATVYGIIDRF 381
 Db 292 KTKGKEDNFTLTSTLKGVTAVODIANRALEDK-----FPGGEHLVY-GLKDG 340
 QY 382 VDIADTSLBGNDKCLATEAISAKKEPEKTKTIPAEVRKTLIPEMD 431
 Db 341 VDLTGYL--NDKT-----KEAVTKADKVISGDIVKPEKPE 375
 RESULT 13
 US-08-396-957A-5
 Sequence 5, Application US/08396957A
 Patent No. 5780041
 GENERAL INFORMATION:
 APPLICANT: SIMPSON, WARREN; SCHWAN, TOM G.
 TITLE OF INVENTION: ANTIGENIC PROTEINS AND
 TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDOFFERI.
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/396,957A
 FILING DATE: 01-MAR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/020,245
 FILING DATE: 19-FEB-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/664,731
 FILING DATE: 05-MAY-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/487,716
 FILING DATE: 05-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: DOROTHY R. AUTH
 REGISTRATION NUMBER: 36,436
 REFERENCE/DOCKET NUMBER: 2026-4018054
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 341
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Borrelia burgdorferi
 STRAIN: Sh-2-82
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:

CELL LINE:
 ORGANISM:
 FEATURE:
 NAME/KEY: p39a
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION: p39a protein sequence
 US-08-396-957A-5
 Query Match 7.8%; Score 179; DB 1; Length 341;
 Best Local Similarity 26.9%; Pred. No. 5.8e-08;
 Matches 101; Conservative 59; Mismatches 123; Indels 92; Gaps 25;
 QY 72 ANKHFGLNMAIVTAGTVNDNSFNQSSWEAIIQALGALTGSGITSVDSSTALECKYS-- 129
 Db 24 SSKIKIXISMLV---DGLVLDKSFSSANVELLRKXKFPENIEVFP--CAISGVSSYV 78
 QY 130 --LANTNKN---VWVLSGFQHGDAFTRWLKIPEKKOLFTEKNII--ILGIDWTDENVY 181
 Db 79 SDLNLRKNSGDLTW-LVGYWLTDA--SLVSSENPKI--SYGIIDPIYGDVQIPEMLI 133
 QY 182 PTGVYILTYKTEBAGLACY--ANASFLAKKFPSPDTRKSAIVIG--GGISPAVTD-FI 236
 Db 134 A-----VFRVBSQAGLAGIAKKSFSGK-----IGFTGKKNIVDAFR 174
 QY 237 AGYLAGIKANNLKNKSDKTKITTDKIEINLFPVDQSTYKRLBOIASK--DKPSTLLAV 294
 Db 175 YGVESGAKYAN-KDIELISEVSNFSFSDVIG-----RTIASKQYSKIDIVYHF 221
 QY 295 AGPLTEI-FSDIIANQND-RYLIGVDTQSLVYTKTKNKFSTSLKNLGYSPVLSDLX 352
 Db 222 AAGLAGIGVIEAANKLGDGYVIGADQDS--YLAPKN-FITSYIKNIGDALYLTGS-Y 277
 QY 353 TKXSNRNLAGFERGKSAATVYIGIKRFPVADIADTSLBGNDKCLATEAISAKKEPEKT 412
 Db 278 IKNNNV-----WESGK--VWQGLRQDVGILPMAN-----EF-EYI 310
 QY 413 KTIPEAEVRKTLIIP 427
 Db 311 KVLERKIVNKEIIVP 325
 RESULT 14
 US-09-182-625F-6
 Sequence 6, Application US/09182625F
 Patent No. 6506892
 GENERAL INFORMATION:
 APPLICANT: Webb, Andrew C.
 APPLICANT: Blazat, Beverly A.
 TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gr
 TITLE OF INVENTION: Regulation
 FILE REFERENCE: BLAZ-101XCI
 CURRENT APPLICATION NUMBER: US/09/182,625F
 CURRENT FILING DATE: 1998-10-29
 PRIOR APPLICATION NUMBER: US 60/063,701
 PRIOR FILING DATE: 1997-10-29
 NUMBER OF SEQ ID NOS: 16
 SEQ ID NO 6
 LENGTH: 166
 TYPE: PRT
 ORGANISM: Mycoplasma fermentans
 FEATURE:
 OTHER INFORMATION: Peptide
 US-09-182-625F-6
 Query Match 7.5%; Score 173; DB 4; Length 166;
 Best Local Similarity 27.3%; Pred. No. 6.6e-08;
 Matches 50; Conservative 37; Mismatches 56; Indels 40; Gaps 8;
 QY 225 GGGISPAVTDPIAGYLAGIKANNLKNKSDKTKI-TTKIKIRINIGFVDQSTYKRLBOIA 283
 Db 2 GGAFAFPGVTTFNBSGAFGLIYNOKH--KSSKIYHTSPVLKDSGF-----TAGKKNMTVI 54

```

Db          28 S-----KXPESTLLAVAGPLTEIFSDIIANQNDRIYILGVDTDQSLVYTKRKF 333
QY          334 TSLILKMGVSVSVSLD-----YTKSNSRNLAGEFPGKSAATVYLGIKDRFVD 383
Db          111 TSVLKHAIKAIVAVETLLDILIEKBEQYPPVVDKXKADKWSHFGQ-----KERWIG 162
QY          384 IAD 386
Db          163 VAE 165

RESULT 15
US-08-396-957A-4
: Sequence 4, Application US/08396957A
: Patent No. 5780041
:
GENERAL INFORMATION:
: APPLICANT: SIMPSON, WARREN, SCHMAN, TOM G.
: TITLE OF INVENTION: ANTIGENIC PROTEINS AND
: TITLE OF INVENTION: GENES ENCODING SAME OF BORRELIA BURGDORFERI.
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & PINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY DISK
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/396,957A
: FILING DATE: 01-MAR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/020,245
: FILING DATE: 19-FEB-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/664,731
: FILING DATE: 05-MAY-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/487,716
: FILING DATE: 05-MAY-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: DOROTHY R. AUTH
: REGISTRATION NUMBER: 36,434
: REFERENCE/DOCKET NUMBER: 2026-4018USA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 758-4800
: TELEFAX: (212) 751-6849
: TELEX: 421792
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 339
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Borrelia burgdorferi
: STRAIN: Sh-2-82
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
: CELL LINE:

```

```

?      ORGANELLE:
?      FEATURE:
?      NAME/KEY:  p39'
?      LOCATION:
?      IDENTIFICATION METHOD:
?      OTHER INFORMATION:  p39', protein
?      OTHER INFORMATION:  sequence.
US-08-396-957A-4

```

Query Match	7.0%;	Score 162;	DB 1;	Length 339;
Best Local Similarity	22.8%;	Pred. No. 2.1e-06;		
Matches	85;	Conservative	69;	Mismatches 135;
				Indels 84;
				Gaps 18;

```

Qy      87 GTVNDNSPFOSSWEBAIQOLGALTGEIITSVDSST-----AELEGKYSLSLANTNNKRWVLG 142
Db      37 GTFDKDSFNFESALNNGVKYKKEBFKEIWLKSSSSSYLSLDEG-----LKGAGSLIWLIG 92

Qy      143 FOHGD-AFTRMUKIPENKQLFTEKNIIILGIDWDTENVIPTRYINULTYKTEBAGLAG 201
Db      93 YRFSVDAKVAALQNPDMKTAI-----IDPIYSNDPIF-ANLVGMTFAREGAELTV 142

Qy      202 YNASTLAKKPFSDDPKTSAYIYGGISPAVTD-FLAGLAGIARWLKNSDKKTKITTD 266
Db      143 YIAAKL-----SKGKIGELGIGELVDARRYEGAGAYAN-----KQIKISTQ 188

Qy      261 KI-----EIMLGFVDQDSTFKERLEQIASKQKPTLLAVAGLPEIFSDIIANQ--NDRYL 314
Db      189 YISSPDLNLAGSVATRMYSDEID-----IHHNADLGIGALIVAKELGSGHYI 238

Qy      315 IGVDTQDSLVYTKTKKKEFTSLIKNLGYSVFSVSLDLYTKKSNRNLAGEFEGKKSATVY 374
Db      239 IGVDEDAQVL--APDNVITISYTKOVGRAL-----NIPT--SNHLKTNTFEGGK--LIN 285

Qy      375 LGIKDFVDIADTSLBGNDKLATATAISSAKKEPEPEKTKTIPAEVAVKTEIPEMPDKOP 434
Db      286 YGLKEGVAVFV-----RNPKMISF-----ELEKEIDNLSKKINKKEIIVP----- 325

Qy      435 DKQGESLDLKITD 447
Db      326 -SNKESYEKFLKE 337

```

Search completed: December 18, 2004, 01:28:35
Job time : 33.9611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:26:40 ; Search time 104.236 Seconds
(without alignments)
1548.256 Million cell updates/sec

Title: US-09-676-249d-2
Perfect score: 2299
Sequence: 1 MKKKIKMKFKLGLVPLS.....KPPDKQESLDTITINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgnt_6/ptodaca/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgnt_6/ptodaca/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgnt_6/ptodaca/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgnt_6/ptodaca/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgnt_6/ptodaca/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgnt_6/ptodaca/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgnt_6/ptodaca/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgnt_6/ptodaca/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgnt_6/ptodaca/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgnt_6/ptodaca/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgnt_6/ptodaca/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgnt_6/ptodaca/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgnt_6/ptodaca/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgnt_6/ptodaca/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgnt_6/ptodaca/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgnt_6/ptodaca/2/pubpaa/US10_PUBCOMB.pep:*
17: /cgnt_6/ptodaca/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgnt_6/ptodaca/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgnt_6/ptodaca/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgnt_6/ptodaca/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	232.5	10.1	350	US-10-474-792-166	Sequence 166, App
2	223	9.7	350	US-09-769-787-132	Sequence 132, App
3	221	9.6	344	US-10-451-337-6	Sequence 6, Appl
4	217.5	9.5	330	US-10-451-337-12	Sequence 12, Appl
5	217.5	9.5	330	US-10-451-337-14	Sequence 14, Appl
6	217.5	9.5	330	US-10-451-337-39	Sequence 39, Appl
7	217.5	9.5	330	US-10-451-337-40	Sequence 40, Appl
8	217.5	9.5	330	US-10-451-337-41	Sequence 41, Appl
9	214.5	9.3	330	US-10-451-337-16	Sequence 16, Appl
10	214.5	9.3	330	US-10-451-337-42	Sequence 42, Appl
11	203	8.8	328	US-09-765-272-8	Sequence 8, Appl
12	188.5	8.2	347	US-10-451-337-8	Sequence 8, Appl
13	188.5	8.2	347	US-10-474-792-180	Sequence 180, App

14	182	7.9	339	US-09-071-035-80	Sequence 80, Appl
15	182	7.9	339	US-10-206-576-80	Sequence 80, Appl
16	182	7.9	361	US-09-071-035-78	Sequence 78, Appl
17	182	7.9	361	US-10-206-576-78	Sequence 78, Appl
18	173	7.5	166	US-09-760-541-6	Sequence 6, Appl
19	169	7.4	357	US-10-282-122A-60484	Sequence 60484, A
20	156	6.8	797	US-10-156-761-10907	Sequence 10907, A
21	150	6.5	223	US-09-769-736-48	Sequence 48, Appl
22	141.5	6.2	889	US-09-952-267-15	Sequence 15, Appl
23	131.5	5.7	1786	US-09-742-096-3	Sequence 3, Appl
24	131.5	5.7	1787	US-10-415-253-2	Sequence 2, Appl
25	130.5	5.7	496	US-10-451-467A-432	Sequence 432, App
26	123.5	5.4	998	US-10-282-122A-70450	Sequence 70450, A
27	123.5	5.4	1031	US-10-282-122A-54611	Sequence 54611, A
28	123.5	5.3	2045	US-10-282-122A-74463	Sequence 74463, A
29	122.5	5.3	2059	US-10-474-792-62	Sequence 62, Appl
30	122.5	5.3	2285	US-09-932-183A-2	Sequence 2, Appl
31	121.5	5.3	852	US-10-282-122A-62892	Sequence 62892, A
32	121.5	5.3	1104	US-09-797-385-4	Sequence 4, Appl
33	120.5	5.2	2457	US-10-282-122A-49854	Sequence 49854, A
34	120	5.2	719	US-10-239-610-2	Sequence 2, Appl
35	119	5.2	776	US-10-238-075-571	Sequence 571, App
36	119	5.2	892	US-09-952-267-5	Sequence 5, Appl
37	119	5.2	902	US-10-437-963-105564	Sequence 105564, A
38	118.5	5.2	1963	US-10-282-122A-73978	Sequence 73978, A
39	118.5	5.2	2872	US-10-282-122A-60137	Sequence 60137, A
40	118	5.1	571	US-10-282-122A-70208	Sequence 70208, A
41	118	5.1	896	US-10-282-122A-60816	Sequence 60816, A
42	118	5.1	1849	US-10-637-544-2	Sequence 2, Appl
43	117.5	5.1	1073	US-10-193-764-45	Sequence 45, Appl
44	117.5	5.1	1079	US-10-193-764-43	Sequence 43, Appl
45	117.5	5.1	6641	US-10-282-122A-70580	Sequence 70580, A

ALIGNMENTS

RESULT 1
US-10-474-792-166

Sequence 166, Application US/10474792

Publication No. US20040236072A1

GENERAL INFORMATION:

APPLICANT: Olmsted, Stephen

APPLICANT: Zagursky, Robert

APPLICANT: Nickbarg, Elliot

APPLICANT: Winter, Louie

TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES

FILE REFERENCE: AM 100399

CURRENT APPLICATION NUMBER: US/10/474, 792

CURRENT FILING DATE: 2003-10-14

NUMBER OF SEQ ID NOS: 674

SOFTWARE: PatentIn version 3.0

SEQ ID NO 166

LENGTH: 350

TYPE: PRT

ORGANISM: Streptococcus pyogenes

US-10-474-792-166

Query Match 10.1%, Score 232.5; DB 17; Length 350;
Best Local Similarity 26.3%; Pred. No. 9e-09;
Matches 113; Conservative 57; Mismatches 153; Indels 107; Gaps 20;

QY	9	KPLGLGLVPLSLATIS-AGCWDKETTKEKSDNQNKTITVDYKISGLVNERKSEIMA	67
DB	4	KPLGLG---LASVAVLSLAACNRG-----A	26
QY	68	AKADANRKHFGALNAIVTAGTVDNSFNQSSWEAIOQLGALTG-GRITVSVDSTAELECK	126
DB	27	SKGASAKDTLKVAMVTDGVDKSFNQSAMELQSGKEMGKQGTGFYFQSTSESE	86
QY	127	YSSLIANTNKVNVLSGQ--HGDAFTRWLKIPEKKOLFTEKNIIILIGIDWTDTENVPTG	184
DB	87	YA---TNLDTAVSGGYLIYGIGA--LKDAILAKAGDNEGVFVIID---DIEGKD	136

QY 185 RYINLYKTEBAGWLAGYANASFLAKKPPSDPTKRSAL-VIGGISPAVTDFIAGYLAGI 243
Db 137 NVASVTFAHDEAAVLAGIAAALAK-----TTTKTGVFGMEGTVITTFEKGFBAGV 187
QY 244 KAMULKNSDKKTKTTTDXIEINLGF-----DVODSTKERLEQIASKQKPSSTLAVAGPL 298
Db 188 KS-----VDDTIOVKVDYAGSFGDAKGTITAAQYAAAD--VYQAAAGTG 233
QY 299 TEIFSDIIANONDR-----YLIGVDTQD--SLVYT-----KTKKKFPTSLKNLGYVSFS 346
Db 234 AGVNEAKAIKESSEADKVVIGVDRQKDEGKITSXGKEANFVLASSIKKGVKAVQL 293
QY 347 VLSDLTYTKSSNRNLAGEFEGKKSATVYLGIKDRFVDIADTSLGNDKRLATEALSEAKK 406
Db 294 INKQVADKK-----FPGGK--TTYV-GLKDGVEIATT-----NVSKAVKAIKEAKA 338
QY 407 EFEKTKTIP 416
Db 339 KIKSGDIKVP 348

RESULT 2

US-09-769-787-132
; Sequence 132, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hanebro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P2129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 132
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-132

Query Match 9.7%; Score 223; DB 10; Length 350;
Best Local Similarity 24.8%; Pred. No. 4.5e-08;

Matches 110; Conservative 63; Mismatches 148; Indels 122; Gaps 22;

QY 1 MKKKIKMKNKFLGLGVPLSAIATIS-AGCWDKETTEKESADNONKQITDVSKISGLVN 59
Db 1 MNKK-----QWTLGLIV-----AVAAVGLAACNRRSSRNAASSSDVYTK----- 39
QY 60 ERKSEIMAAKADANKFGLMAIYTAGTVNDNSFNQSWALIQDLA---LTGEITSV 116
Db 40 -----AAIVDTGTGVDDKSFNOSAMEGLDAMGKEHNI,SKONGFTY 79
QY 117 DNSTAELE-----GKYSSLANTKNVWVLSFGQDAFTWMLKIPENKQIFTEKNI 167
Db 80 FQSTSEADYANNLOOAGSTVILFGV-----GFLANNA-----VKDAKERTDILNY 125
QY 168 IILGIDWTDENVIPTRGYINLYKTEBAGWLAGYANASFLAKKPPSDPTKRSALVIGGG 227
Db 126 VLIDVIDIDQKGV-----ASVTFADNSGYLAGYAAAK-----TTTKQVGFVG 170
QY 228 I-SPAVDFIAGYLAGIYKAMULKNSDKKTKTTTDXIEINLGFVDVODSTKERLEQIASKD 286
Db 171 IESSEVISRFEAGFPAAGV-----ASVDPISIKQVVDYAG-SFG-DAKAGKTIAAQYAAAD 223
QY 287 KPSITLAAVAGPL-TEIFSDIIANONDR-----YLIGVDTQD--SLVYT-----KTKKNKF 333
Db 224 ---IVQVAGGTGAGVFAEAKSLNESRPEKNEKVVIGVDRQDEAEKITSKQKESNFVL 280

QY 334 TSLIKNLGYVSFVLSDLTYTKSSNRNLAGEFEGKKSATVYLGIKDRFVDIADTSLGND 393
Db 281 VSTLKQVGTIVKDI-----SNKARCEFPQGQ---VVIYSIKDXKGVDLAVTNLSSEG 329
QY 394 KRLATEALSEAKKEFEKTKTIP 416
Db 330 KK-----AVEDAKAKILDGSVKVP 348

RESULT 3

US-10-451-337-6
; Sequence 6, Application US/10451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denise
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Joee
; APPLICANT: RIOUX, Stephane
; APPLICANT: RHEAULT, Patrick
; TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
; TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS
; FILE REFERENCE: 12806-24RPT
; CURRENT APPLICATION NUMBER: US/10/451,337
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 344
; TYPE: PRT
; ORGANISM: S. pyogenes
US-10-451-337-6

Query Match 9.6%; Score 221; DB 15; Length 344;
Best Local Similarity 26.4%; Pred. No. 6.1e-08;

Matches 115; Conservative 56; Mismatches 141; Indels 124; Gaps 23;

QY 9 KFLGLGVPLSAIATIS-AGCWDKETTEKESADNONKQITDVSKISGLVNERKSEIMA 67
Db 3 KFLGLG-----LASVAVLSLAACNKG-----A 25
QY 68 AKADANKFGLMAIYTAGTVNDNSFNQSWAL---QDLGALTGGEITSVDSSTAELE 124
Db 26 SKGASGKTDLKVMVMTDTGGVDDKSFNOSAMEGQSWGKEWLOKQ---TGFDYFQSTSE 82
QY 125 GKYSSSLANTKNVWVLSGFQ--HGDATPRMLK-----IPENKQIFTEKNIITIIIGIDWTD 176
Db 83 SEVA---TWLDTAVSGGYOLIVGIFA--LKDAIARAAGDGVKFIYIIDILIEGKD-- 133
QY 177 TENYIPTRGYINLYKTEBAGWLAGYANASFLAKKPPSDPTKRSALVIGGISPAVTDFI 236
Db 134 --NV-----ASVTFADHEAAVLAGIAAALAK-----TTTKTGVFGMEGTVITRF 176
QY 237 AGYLAGIKAMULKNSDKKTKTTTDXIEINLGF-----DVODSTKERLEQIASKQKPSSTL 291
Db 177 EGFPAGVKS-----VDDTIOVKVDYAGSFGDAKGTITAAQYAAAD--VIY 222
QY 292 LAVAGPLTEIFSDIIANON--DR-YLIGVDTQD--SLVYT-----KTKKKFPTSLKNL 340
Db 223 QAAAGTGAIVENAIIKESSEADKVVIGVDRQKDEGKITSXGKEANFVLASSIKKGV 282
QY 341 GYSVFSVLSDLTYTKSSNRNLAGEFEGKKSATVYLGIKDRFVDIADTSLGNDKRLATEA 400
Db 283 GKAVQLINKQVADK-----FPGGK--TTYV-GLKDGVEIATT-----NVSKAVKA 326
QY 401 ISEAKKEFEKTKTIP 416
Db 327 IKEAKAKIKSGDIKVP 342


```

RESULT 4
US-10-451-337-12
; Sequence 12, Application US/10451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denis
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Josee
; APPLICANT: RIOUX, Stephane
; APPLICANT: RHEAULT, Patrick
; TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
; FILE REFERENCE: 12806-24PCT
; CURRENT APPLICATION NUMBER: US/10/451,337
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 330
; TYPE: PRT
; ORGANISM: S. pyogenes
US-10-451-337-12

Query Match          9.5%; Score 217.5; DB 15; Length 330;
Best Local Similarity 28.3%; Pred. No. 1e-07;
Matches 102; Conservative 47; Mismatches 142; Indels 69; Gaps 17

QY      67 AAKADANHFGLNMAIVTAGTVNNNSFNOSWEALIQOLGALTG-CETISVSVSTAELEG 125
       |||::||::||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      19 ASKGASAKTDLKVMAMVTDTGGVDKSPFQSAMEGIQSGKEMGLQKGFYPOSTSES 78

QY      126 KYSLIANTNKWVVVVSFGFO--HGDAFTBWLKIPENKKOLFTEKNIIILGIDMTDFENVIPT 183
       :::::--TNLDTVASGCGYOLITGIGA--LKAIKAAGADNGCAVFVIID-----DIEEK 128
DB      79 EYA-----TNLDTVASGCGYOLITGIGA--LKAIKAAGADNGCAVFVIID-----DIEEK 128

QY      184 GRYNLYTKTBEGAGVLAGYANAFLAKKPPSDPYRSAL-VIGGISPAVTFPIAGYIAG 242
       :::::--DNVASVTPADHBAAYLAGIAAAK-----TTTKKTGVFGMGEGTVITRFKGFPEAG 179
DB      129 DNVASVTPADHBAAYLAGIAAAK-----TTTKKTGVFGMGEGTVITRFKGFPEAG 179

QY      243 IKANMLKNSDKKTKTTTDKIEINLGF-----DVQOTSTERLEQLASHCKRSTLLAVAGP 297
       :::::--VDDPTIOVKVDYASFGDPAAGKTIAMAQAAGAD--VIYQAAAGT 225
DB      180 VKS-----VDDPTIOVKVDYASFGDPAAGKTIAMAQAAGAD--VIYQAAAGT 225

QY      298 LTELPSDIIANONDR-----YLIGVDPDQ--SLVYT-----KTXKKFFTSILKNIGSYVF 345
       :::::--:::--:::--:::--:::--:::--:::--:::--:::--:::--:::--:::
DB      226 GAGVPNEAKAINERKSEADKVAVIGVDRDQKBEGRKITSODGEANFVLASSIKVEGKAVQ 285

QY      346 SVLSLDLYTKKSNRNLAGFEFGKKSATYYLGIKDRFVIDIADTSLBGNDKQLATEAISXAK 405
       ::::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
DB      286 LINKOVADKK-----FPGGK--TTYV-GKDGGEIATL---NVSKEAVKAIKEAK 330

RESULT 5
US-10-451-337-14
; Sequence 14, Application US/10451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denis
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Josee
; APPLICANT: RIOUX, Stephane
; APPLICANT: RHEAULT, Patrick
; TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
; FILE REFERENCE: 12806-24PCT
; CURRENT APPLICATION NUMBER: US/10/451,337
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; PRIOR FILING DATE: 2000-12-21

```

```

: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 330
: TYPE: PRT
: ORGANISM: S. pyogenes
US-10-451-337-14

Query Match          9.5%, Score 217.5; DB 15; Length 330;
Best Local Similarity 28.3%, Pred. No. 1e-07;
Matches 102; Conservative 47; Mismatches 142; Indels 69; Gaps 17

Dh 67 AAKADANHGFLNMAIYAGCTVNDNSFNQGSWEAIIQQLGALTG-GEITSVDSSTAELRG 125
19 ASKSGAGSKTDLKYAMVYDTGCVDDKSNQGSABELOSQMGKEMGLQKGTGPFYPOSTSS 78

Dh 126 KYSSLANTNKNVWLSGFO--HGDAFTRWLKI PENKOLFTEKNIIILGIDWTDTEENVLP 183
79 EYA-----TNLDPTAVSGGYQLIYGIGA--LKDALAKAGDNEGVFYIID-----DIIEGK 128

Dh 184 GRYINLTYKTEAGMLAGYANASFLAKFPBDPTKSAI-VIGGISAPVVDPIAGIYLAG 242
129 DNVA5VTFADHEAAYLAGIAAAK-----TTKTKTVGFVGGEGEYVITREFEKGFEAG 179

Dh 243 IKANNLNKSDKKTITTDKIEINIGF-----DVQDTSKERLEQJASDKPSTLLA5AGP 297
180 VKS-----VDDTIQVKVDYAGSPEDBAKAGKTIIAAQYAAAGD--VIYQAGGT 225

Dh 298 LTFEIPSDIIANQNR-----YLIGVPTDQ--SLVYT---KTKNKFPTSILKNLQSYVF 345
226 GAGVFNEAKKAIENKRSKADKVMVIGVDRDQKDEGKYTSKDGKEANFVLASSIKYGVKAVQ 285

Dh 346 SVLSDLTYTKNSNRNLGAFEEGKKSATYYLGIKDRFVDIADTSLGNDKLLATEAISEAK 405
286 LINKQVADKK-----FPGGK--TTYV-GLKGGVGIATY---NVSKEAVKAIKEAK 330

RESULT 6
US-10-451-337-39
: Sequence 39, Application US/10451337
: Publication No. US20040097706A1
: GENERAL INFORMATION:
: APPLICANT: SHIRE BIOCHEM INC.
: APPLICANT: MARTIN, Denis
: APPLICANT: BRODEUR, Bernard R.
: APPLICANT: HAMEL, Josee
: APPLICANT: RIOUX, Stephane
: APPLICANT: RHEAULT, Patrick
: TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
: FILE OF INVENTION: CORRESPONDING DNA FRAGMENTS
: FILE REFERENCE: 12806-24PCT
: CURRENT APPLICATION NUMBER: US/10/451.337
: CURRENT FILING DATE: 2003-11-18
: PRIOR APPLICATION NUMBER: US 60/256,940
: PRIOR FILING DATE: 2000-12-21
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 39
: LENGTH: 330
: TYPE: PRT
: ORGANISM: S. pyogenes
US-10-451-337-39

Query Match          9.5%, Score 217.5; DB 15; Length 330;
Best Local Similarity 28.3%, Pred. No. 1e-07;
Matches 102; Conservative 47; Mismatches 142; Indels 69; Gaps 17

Dh 67 AAKADANHGFLNMAIYAGCTVNDNSFNQGSWEAIIQQLGALTG-GEITSVDSSTAELRG 125
19 ASKSGAGSKTDLKYAMVYDTGCVDDKSNQGSABELOSQMGKEMGLQKGTGPFYPOSTSS 78

Dh 126 KYSSLANTNKNVWLSGFO--HGDAFTRWLKI PENKOLFTEKNIIILGIDWTDTEENVLP 183
79 EYA-----TNLDPTAVSGGYQLIYGIGA--LKDALAKAGDNEGVFYIID-----DIIEGK 128

Dh 184 GRYINLTYKTEAGMLAGYANASFLAKFPBDPTKSAI-VIGGISAPVVDPIAGIYLAG 242
129 DNVA5VTFADHEAAYLAGIAAAK-----TTKTKTVGFVGGEGEYVITREFEKGFEAG 179

Dh 243 IKANNLNKSDKKTITTDKIEINIGF-----DVQDTSKERLEQJASDKPSTLLA5AGP 297
180 VKS-----VDDTIQVKVDYAGSPEDBAKAGKTIIAAQYAAAGD--VIYQAGGT 225

Dh 298 LTFEIPSDIIANQNR-----YLIGVPTDQ--SLVYT---KTKNKFPTSILKNLQSYVF 345
226 GAGVFNEAKKAIENKRSKADKVMVIGVDRDQKDEGKYTSKDGKEANFVLASSIKYGVKAVQ 285

Dh 346 SVLSDLTYTKNSNRNLGAFEEGKKSATYYLGIKDRFVDIADTSLGNDKLLATEAISEAK 405
286 LINKQVADKK-----FPGGK--TTYV-GLKGGVGIATY---NVSKEAVKAIKEAK 330

```

Db 79 EYA---TNDLTA VSGGYQLIYGIFA--LKDAIAKAGDNEGVKFIID---DIIIEGK 128

Qy 184 GRVYINLTYTEAGWLAGYANASFLAKKPSDPTKRSAL-VIGGISPAAVDTFLAGYLAG 242

Db 129 DNVA SVTFADHEAAYLAGIAAIAK-----TTKTKTVGFGMEGVITIRFEKGFEAG 179

Qy 243 IKAWNLNKSDKTKITTTDKIEINIGF---DVODSTKERLEQIASKDKPSTLLA VAGP 297

Db 180 VKS-----VDDTIQVKVDYAGSFGDAKAGKTI AAQYAAAGAD--VIYQAAAGGT 225

Qy 298 LTFEFSDIANQNR-----YLIQVDTDO--SLVYT---KTGKKEFTSLKXLYGSVF 345

Db 226 GAGVFNEKKAINEKRSADKVVIGVDRDQDEGKRTYSKDGKEANFVLASSIKEVGA VQ 285

Qy 346 SVLSDLYTKKSNSHNLGAFEFKKSATVYLGIKDRFYDIADTSLGNDKCLATEAISEAK 405

Db 286 LINKQVADKK-----FPGGK--TTYV-GLKDGGEIATTT---NVSKEAVAIKAEAK 330

RESULT 7

US-10-451-337-40

Sequence 40, Application US/10451337

Publication No. US20040097706A1

GENERAL INFORMATION:

APPLICANT: SHIRE BIOCHEM INC.

APPLICANT: MARTIN, Denis

APPLICANT: BRODEUR, Bernard R.

APPLICANT: HAMEL, Josee

APPLICANT: RIOUX, Stephane

APPLICANT: RHEAULT, Patrick

TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND

TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS

FILE REFERENCE: 12806-24PCT

CURRENT APPLICATION NUMBER: US/10/451.337

CURRENT FILING DATE: 2003-11-18

PRIOR APPLICATION NUMBER: US 60/256,940

PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 42

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 40

LENGTH: 330

TYPE: PRT

ORGANISM: S. pyogenes

US-10-451-337-40

Query Match 9.5%; Score 217.5; DB 15; Length 330;

Best Local Similarity 28.3%; Pred. No. 1e-07;

Matches 102; Conservative 47; Mismatches 142; Indels 69; Gaps 17;

Qy 67 AAKADANKHFGLNMAIYTAGGTVDNSFNQSSWEAIIQQLGALTG-GEITSVDSSTALEG 125

Db 19 ASKGGASGKTDLKYAMVTDGTGVDKSFNQSAMEGLQSWGKEMGLQKGTGFDYFQSTSES 78

Qy 126 KYSSLANTNNKVVWLSGFO--HGDAFTRWLKIPENKOLFTEKNIIIGIDWTDENVIPT 183

Db 79 EYA---TNDLTA VSGGYQLIYGIFA--LKDAIAKAGDNEGVKFIID---DIIIEGK 128

Qy 184 GRVYINLTYTEAGWLAGYANASFLAKKPSDPTKRSAL-VIGGISPAAVDTFLAGYLAG 242

Db 129 DNVA SVTFADHEAAYLAGIAAIAK-----TTKTKTVGFGMEGVITIRFEKGFEAG 179

Qy 243 IKAWNLNKSDKTKITTTDKIEINIGF---DVODSTKERLEQIASKDKPSTLLA VAGP 297

Db 180 VKS-----VDDTIQVKVDYAGSFGDAKAGKTI AAQYAAAGAD--VIYQAAAGGT 225

Qy 298 LTFEFSDIANQNR-----YLIQVDTDO--SLVYT---KTGKKEFTSLKXLYGSVF 345

Db 226 GAGVFNEKKAINEKRSADKVVIGVDRDQDEGKRTYSKDGKEANFVLASSIKEVGA VQ 285

Qy 346 SVLSDLYTKKSNSHNLGAFEFKKSATVYLGIKDRFYDIADTSLGNDKCLATEAISEAK 405

Db 286 LINKQVADKK-----FPGGK--TTYV-GLKDGGEIATTT---NVSKEAVAIKAEAK 330

RESULT 8

US-10-451-337-41

Sequence 41, Application US/10451337

Publication No. US20040097706A1

GENERAL INFORMATION:

APPLICANT: SHIRE BIOCHEM INC.

APPLICANT: MARTIN, Denis

APPLICANT: BRODEUR, Bernard R.

APPLICANT: HAMEL, Josee

APPLICANT: RIOUX, Stephane

APPLICANT: RHEAULT, Patrick

TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND

TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS

FILE REFERENCE: 12806-24PCT

CURRENT APPLICATION NUMBER: US/10/451.337

CURRENT FILING DATE: 2003-11-18

PRIOR APPLICATION NUMBER: US 60/256,940

PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 42

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 41

LENGTH: 330

TYPE: PRT

ORGANISM: S. pyogenes

US-10-451-337-41

Query Match 9.5%; Score 217.5; DB 15; Length 330;

Best Local Similarity 28.3%; Pred. No. 1e-07;

Matches 102; Conservative 47; Mismatches 142; Indels 69; Gaps 17;

Qy 67 AAKADANKHFGLNMAIYTAGGTVDNSFNQSSWEAIIQQLGALTG-GEITSVDSSTALEG 125

Db 19 ASKGGASGKTDLKYAMVTDGTGVDKSFNQSAMEGLQSWGKEMGLQKGTGFDYFQSTSES 78

Qy 126 KYSSLANTNNKVVWLSGFO--HGDAFTRWLKIPENKOLFTEKNIIIGIDWTDENVIPT 183

Db 79 EYA---TNDLTA VSGGYQLIYGIFA--LKDAIAKAGDNEGVKFIID---DIIIEGK 128

Qy 243 IKAWNLNKSDKTKITTTDKIEINIGF---DVODSTKERLEQIASKDKPSTLLA VAGP 297

Db 180 VKS-----VDDTIQVKVDYAGSFGDAKAGKTI AAQYAAAGAD--VIYQAAAGGT 225

Qy 298 LTFEFSDIANQNR-----YLIQVDTDO--SLVYT---KTGKKEFTSLKXLYGSVF 345

Db 226 GAGVFNEKKAINEKRSADKVVIGVDRDQDEGKRTYSKDGKEANFVLASSIKEVGA VQ 285

Qy 346 SVLSDLYTKKSNSHNLGAFEFKKSATVYLGIKDRFYDIADTSLGNDKCLATEAISEAK 405

Db 286 LINKQVADKK-----FPGGK--TTYV-GLKDGGEIATTT---NVSKEAVAIKAEAK 330

RESULT 9

US-10-451-337-16

Sequence 16, Application US/10451337

Publication No. US20040097706A1

GENERAL INFORMATION:

APPLICANT: SHIRE BIOCHEM INC.

APPLICANT: MARTIN, Denis

APPLICANT: BRODEUR, Bernard R.

APPLICANT: HAMEL, Josee

APPLICANT: RIOUX, Stephane

APPLICANT: RHEAULT, Patrick

TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND

TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS

FILE REFERENCE: 12806-24PCT

CURRENT APPLICATION NUMBER: US/10/451.337

CURRENT FILING DATE: 2003-11-18

PRIOR APPLICATION NUMBER: US 60/256,940

PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 330
TYPE: PRT
ORGANISM: S. pyogenes
US-10-451-337-16

Query Match 9.3%; Score 214.5; DB 15; Length 330;
Best Local Similarity 28.3%; Pred. No. 1.7e-07;
Matches 102; Conservative 46; Mismatches 143; Indels 69; Gaps 17;

QY 67 AAKADAKHFGALMAITVAGTVDNSFNOSWEAIOQLGALTG-GETSVDSSTAELEG 125
DB 19 ASKGAAGKTDLAKAMTDTGCVDDKSPNOSAMEGLSWMGEMLOKGTGTFDYQSTSES 78
QY 126 KYSLANTNKVWVLSGFGQ--HGDAFTRWLKIPEKNIILIGIDWTDTEENVIPT 183
DB 79 EVA-----TNDTAVSGGYQLIYGIGA--LKDAIAKAGDNEGVKFIID----DIIEGK 128
QY 184 GRYINLTYTEEAGWLAGVNASFLAKKPSDPTKSAI-VIGGISPATVDFIAGYIAG 242
DB 129 DNVAATVPADHEAAVYLAGIAAK-----TTKTVGVFGMEGTVITRFKGEAG 179
QY 243 IKAMNLKNSDKTKITTDKIEINLGF-----DVQDSTKERLEQIASKDKPSTLLAAGP 297
DB 180 VKS-----VDDTIQVKVDYAGSFGDAKAGKTIAAQYAAAGAD--VIYQAAAGT 225
QY 298 LTEIFSDIINQDR-----YLIGVDTDQ--SLVYT---KTKNKFSTILKNLGYSVF 345
DB 226 GAGVFNKAIAINEKRSBADKVVWIGVDRDQDEGKYTSKDGKEANFVLASSIKVGRVQ 285
QY 346 SVLSDLTKKSNRNLAGFPERGKSAIVYLGIDRFPVDIADTSLEGNDKLATPAISEAK 405
DB 286 LINKQVADK-----FPGK--TTVY-GLKDGVEIAT--NVSKEAVAKIAEAK 330

RESULT 10
US-10-451-337-42
Sequence 42, Application US/10451337
Publication No. US20040097706A1
GENERAL INFORMATION:
APPLICANT: SHIRE BIOCHEM INC.
APPLICANT: MARTIN, Denis
APPLICANT: BRODEUR, Bernard R.
APPLICANT: HAMEL, Josee
APPLICANT: RIOUX, Stephane
APPLICANT: RHEAULT, Patrick
TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS
FILE REFERENCE: 12806-24PCT
CURRENT APPLICATION NUMBER: US/10/451.337
CURRENT FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: US 60/256,940
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 42
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 330
TYPE: PRT
ORGANISM: Mouse
US-10-451-337-42

Query Match 9.3%; Score 214.5; DB 15; Length 330;
Best Local Similarity 28.3%; Pred. No. 1.7e-07;
Matches 102; Conservative 46; Mismatches 143; Indels 69; Gaps 17;

QY 67 AAKADAKHFGALMAITVAGTVDNSFNOSWEAIOQLGALTG-GETSVDSSTAELEG 125
DB 19 ASKGAAGKTDLAKAMTDTGCVDDKSPNOSAMEGLSWMGEMLOKGTGTFDYQSTSES 78
QY 126 KYSLANTNKVWVLSGFGQ--HGDAFTRWLKIPEKNIILIGIDWTDTEENVIPT 183

DB 79 EVA-----TNDTAVSGGYQLIYGIGA--LKDAIAKAGDNEGVKFIID----DIIEGK 128
QY 184 GRYINLTYTEEAGWLAGVNASFLAKKPSDPTKSAI-VIGGISPATVDFIAGYIAG 242
DB 129 DNVAATVPADHEAAVYLAGIAAK-----TTKTVGVFGMEGTVITRFKGEAG 179
QY 243 IKAMNLKNSDKTKITTDKIEINLGF-----DVQDSTKERLEQIASKDKPSTLLAAGP 297
DB 180 VKS-----VDDTIQVKVDYAGSFGDAKAGKTIAAQYAAAGAD--VIYQAAAGT 225
QY 298 LTEIFSDIINQDR-----YLIGVDTDQ--SLVYT---KTKNKFSTILKNLGYSVF 345
DB 226 GAGVFNKAIAINEKRSBADKVVWIGVDRDQDEGKYTSKDGKEANFVLASSIKVGRVQ 285
QY 346 SVLSDLTKKSNRNLAGFPERGKSAIVYLGIDRFPVDIADTSLEGNDKLATPAISEAK 405
DB 286 LINKQVADK-----FPGK--TTVY-GLKDGVEIAT--NVSKEAVAKIAEAK 330

RESULT 11
US-09-765-272-8
Sequence 8, Application US/09765272
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765.272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-765-272-8

Query Match 8.8%; Score 203; DB 9; Length 328;
Best Local Similarity 25.7%; Pred. No. 1.2e-06;
Matches 98; Conservative 56; Mismatches 144; Indels 84; Gaps 19;

QY 61 KRSRIMAKKDHGFGALMAITVAGTVDNSFNOSWEAIOQLGALTG-GETSVDSSTAELEG 117
DB 3 RSRNNAASSDVK---TKAIVITDTGCVDDKSPNOSAMEGLSWMGEMLOKGTGTFDYQSTSES 78
QY 118 SSTAEL-----GKYSANTNKVWVLSGFGHGDFTRWLKIPEKNIILIGIDWTDTEENVIPT 168

```
Db 59 QSTSEADYANNLQQAAGSYNLI PGV-----GFALNNA-----VDAKKEHTDLNLYV 104
Qy 169 ILIGDITDTEENVLPTRGRYINLTYTEAGWLAGYANASFLAKFPSPDPTKSAIVIGGGI 228
Db 105 LIDVIDIQDQKV-----ASVTPADNESGYLAGVAANK-----TTTKQVGFVGGI 149
Qy 229 -SPAVIDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASDK 287
Db 150 ESEVISRFEAGFKAGV-----ASVDPSIKVQVDYAG-SFG-DAAKGKTIAAQYAAGAD- 201
Qy 288 PSTLLAAGPL-TEIFSDIIANQDR-----YLIGVDTDSL--VYT-----KTKNKEFT 334
Db 202 --IVYQVAGTGAGVFAAKSLNESRPENEKRWVIGVDRDQABGKTSKDGKSNFVLV 259
Qy 335 SILKMLGYSPVSLDLYTKKSNRNLAGFPGKKSATVYLGIDRFVADIADTSLEGNDK 394
Db 260 STLKQVGTIVDI-----SNKAEREPFGCG--VIVYSLDKGVDLAVTMLSEBG 308
Qy 395 KLATEAISEAKKEFEKTKTIP 416
Db 309 K---AVEDAKAKIILDSGVKP 326
```

RESULT 12

```
US-10-451-337-8
; Sequence 8, Application US/10451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denis
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Josee
; APPLICANT: RIOUX, Stephane
; APPLICANT: RHEAULT, Patrick
; TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
; TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS
; FILE REFERENCE: 12806-24PCT
; CURRENT APPLICATION NUMBER: US/10/451,337
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 347
; TYPE: PRT
; ORGANISM: S. pyogenes
US-10-451-337-8
```

Query Match 8.2%; Score 188.5; DB 15; Length 347;
Best Local Similarity 25.7%; Pred. No. 1.5e-05;
Matches 113; Conservative 59; Mismatches 150; Indels 117; Gaps 25;

```
Qy 1 MKKKIKNNKFLGGLVPEPLSAIATISAGCWDKETTKEKSADNONKQITDVSKISGLVNE 60
Db 1 MNKKV-----MSLGLV--STALFTL-GGC-----TNSAKQTTDNS----- 33
Qy 61 RKSEIMAKADANKHFGIINAIYTAGGTVDNDSFNQSGMEALQQLG---ALTGE---I 113
Db 34 -----LKIAMITNQIGIDKSFNQSAWEGLOQWKGKENTLEKKGXYDF 76
Qy 114 TSVDSS--TAELEKYSLSLANTKNVWVLSGFQHGDAFTR-WLKI PENKQLFTEKNIIIL 170
Db 77 QSANESFPTNLE---SAVTNGYNLVFGI-GFPLHDAVEKVAANNPDNH--FAIVDVIK 130
Qy 171 GIDWTDTEENVLPTRGRYINLTYTEBAGWLAGYANASFLAKFPSPDPTKSAIVIGGGISP 230
Db 131 G-----QKNV-----ASITPSDHEAAVLAVGVAANK-----TTTKQVGFV-GGMEGD 171
Qy 231 AVTDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASDKRST 290
Db 172 VVKRFEKGFEGAGVSV-----DTIKVRAVYAGSPADAAGKGTIAAQAABGADV 221
```

```
Qy 291 LL-AVAGPLTEIFSDIIANQDR-----YLIGVDTDSL--VYT-----KTKNKEFTSIL 337
Db 222 IYHAAAGTGAGVFSEAKSINERKEEDKRWVIGVDRDQABGKTTTKDGSANFVLTSSI 281
Qy 338 KNLGYSFVSLDLYTKKSNRNLAGFPGKKSATVYLGIDRFVADIADTSLEGNDKLA 397
Db 282 KEVGKALVKYA---VYTSBD-----QFPGGQITTF-GLKEGVSILTTDALTDQTK-- 328
Qy 398 TEAISEAKKEFEKTKTIP 416
Db 329 --AIEAKKAIIBGTITVP 345
```

RESULT 13

```
US-10-474-792-180
; Sequence 180, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Louie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100398
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 180
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-180
```

Query Match 8.2%; Score 188.5; DB 17; Length 347;
Best Local Similarity 25.7%; Pred. No. 1.5e-05;
Matches 113; Conservative 59; Mismatches 150; Indels 117; Gaps 25;

```
Qy 1 MKKKIKNNKFLGGLVPEPLSAIATISAGCWDKETTKEKSADNONKQITDVSKISGLVNE 60
Db 1 MNKKV-----MSLGLV--STALFTL-GGC-----TNSAKQTTDNS----- 33
Qy 61 RKSEIMAKADANKHFGIINAIYTAGGTVDNDSFNQSGMEALQQLG---ALTGE---I 113
Db 34 -----LKIAMITNQIGIDKSFNQSAWEGLOQWKGKENTLEKKGXYDF 76
Qy 114 TSVDSS--TAELEKYSLSLANTKNVWVLSGFQHGDAFTR-WLKI PENKQLFTEKNIIIL 170
Db 77 QSANESFPTNLE---SAVTNGYNLVFGI-GFPLHDAVEKVAANNPDNH--FAIVDVIK 130
Qy 171 GIDWTDTEENVLPTRGRYINLTYTEBAGWLAGYANASFLAKFPSPDPTKSAIVIGGGISP 230
Db 131 G-----QKNV-----ASITPSDHEAAVLAVGVAANK-----TTTKQVGFV-GGMEGD 171
Qy 231 AVTDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASDKRST 290
Db 172 VVKRFEKGFEGAGVSV-----DTIKVRAVYAGSPADAAGKGTIAAQAABGADV 221
Qy 291 LL-AVAGPLTEIFSDIIANQDR-----YLIGVDTDSL--VYT-----KTKNKEFTSIL 337
Db 222 IYHAAAGTGAGVFSEAKSINERKEEDKRWVIGVDRDQABGKTTTKDGSANFVLTSSI 281
Qy 338 KNLGYSFVSLDLYTKKSNRNLAGFPGKKSATVYLGIDRFVADIADTSLEGNDKLA 397
Db 282 KEVGKALVKYA---VYTSBD-----QFPGGQITTF-GLKEGVSILTTDALTDQTK-- 328
Qy 398 TEAISEAKKEFEKTKTIP 416
Db 329 --AIEAKKAIIBGTITVP 345
```

RESULT 14

US-09-071-035-80
Sequence 80, Application US/09071035
Publication No. US20020045737A1
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-80
Query Match 7.9%; Score 182; DB 9; Length 339;
Best Local Similarity 25.4%; Pred. No. 4.4e-05;
Matches 104; Conservative 51; Mismatches 139; Indels 116; Gaps 20;
QY 61 RKSEIMAKADANKHFGILNMAIVTAGGVNDNSFNQSSWEAIOQLG--LTGE----- 112
DB 7 KTAESGGGKGA---AHSAVIITDTGGVDKSPFQSGMEGLQAMGKEHDLPEGSKGYAY 62
QY 113 ITSVDSS--TALEGGKSSSLANTKNVWVLSGFQHGDAFTRWLKIPEKQLFTEKNIIIL 170
DB 63 IQSNDADYTTNIDQAVSSKFNITFGI---GYLLKDA-----ISSAADNPPTNPFVLI 112
QY 171 GIDWTDENVIPTRGIYINLYKTEBAGWLAGYANASFLAKKFPSPDPTKRSATVIGGIGISP 230
DB 113 DDQIDGKKV-----VSATFRDNEAAYLAGVAAA-----NETKTNKVGFGGEG 157
QY 231 AVTD-FLAGYLAGIKAMNLKNSDKKTKITTDKIEINLGFVDVDTSTKERLEQIASKDKPS 289
DB 158 VVIDRFQAGFEKGV-ADAKELGKEITVDT-----KYAASFADPA 196
QY 290 TLLAVAGPLTEIFSDII-----ANQNR-YLIGVDTDSL--VYT 326
DB 197 KKKLAAMVONGVILIFHASGATGQGVFOEAKDLNESGSDKVMWIGVDRDQADGKY- 255
QY 327 KTK-----NKEPFTSLKMLGYSVSVSLDLYTKKSNRHLAGFEFGKKSATVYIGIKDRF 381
DB 256 KTKGKKEKNTFLITSLKGVTAVDIARALEDK-----FPGSEHLVY-GLKGG 304
QY 382 VDIADTSLGNDKCLATPAISEAKKEFEKTKTTPAEVVRKLTLEIPEMPD 431
DB 305 VDLTDTGYL--NDKT-----KEAVTAKDKVIGSDVVKPEKPE 339

RESULT 15
US-10-206-576-80
Sequence 80, Application US/10206576
Publication No. US20030017495A1
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-10-206-576-80
Query Match 7.9%; Score 182; DB 14; Length 339;
Best Local Similarity 25.4%; Pred. No. 4.4e-05;
Matches 104; Conservative 51; Mismatches 139; Indels 116; Gaps 20;
QY 61 RKSEIMAKADANKHFGILNMAIVTAGGVNDNSFNQSSWEAIOQLG--LTGE----- 112
DB 7 KTAESGGGKGA---AHSAVIITDTGGVDKSPFQSGMEGLQAMGKEHDLPEGSKGYAY 62
QY 113 ITSVDSS--TALEGGKSSSLANTKNVWVLSGFQHGDAFTRWLKIPEKQLFTEKNIIIL 170
DB 63 IQSNDADYTTNIDQAVSSKFNITFGI---GYLLKDA-----ISSAADNPPTNPFVLI 112
QY 171 GIDWTDENVIPTRGIYINLYKTEBAGWLAGYANASFLAKKFPSPDPTKRSATVIGGIGISP 230
DB 113 DDQIDGKKV-----VSATFRDNEAAYLAGVAAA-----NETKTNKVGFGGEG 157
QY 231 AVTD-FLAGYLAGIKAMNLKNSDKKTKITTDKIEINLGFVDVDTSTKERLEQIASKDKPS 289
DB 158 VVIDRFQAGFEKGV-ADAKELGKEITVDT-----KYAASFADPA 196
QY 290 TLLAVAGPLTEIFSDII-----ANQNR-YLIGVDTDSL--VYT 326
DB 197 KKKLAAMVONGVILIFHASGATGQGVFOEAKDLNESGSDKVMWIGVDRDQADGKY- 255

OY 327 KTK-----NKFPTSLKXNLGYVFSVLSDLVYTKKSNRNLAGFERGKKSATVYLGIDRF 381
Db 256 KTKDGEKEDNFTLTSTLKGVTAVODIANRALEDK-----FPGGEHLVY-GLKXDG 304
OY 382 VDIADTSLGNDKXLATEAISEAKKEPEKTKTIPAEFERKTLRIPEWPD 431
Db 305 VDLTIDGYL--NDKT-----KEAVKTAKDVKVISGDVKVPEKPE 339

Search completed: December 18, 2004, 01:44:02
Job time : 106.236 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:15:13 ; Search time 27.349 Seconds
(without alignments)
1586.669 Million cell updates/sec

Title: US-09-676-249D-2

Perfect score: 2299
Sequence: 1 MKKKIKWKKFGLGLVFPPLS.....KQPDKQOESLDKLTIDINKN 451

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554.5	24.1	461	2 B90555	ABC transporter xy
2	223	9.7	350	2 G95097	lipoprotein (impor
3	219	9.5	374	2 G97965	conserved hypotet
4	207	9.0	350	2 F86804	basic membrane pro
5	195.5	8.5	357	2 D96986	probable lipoprote
6	187.5	8.2	516	2 C82946	hypothetical prote
7	185.5	8.1	350	1 C70009	ABC transporter (l
8	182	7.9	353	1 H71340	membrane lipoprote
9	181.5	7.9	359	2 F72418	basic membrane pro
10	178	7.7	341	2 F70147	basic membrane lip
11	174	7.6	357	2 AH1610	CD4+ T cell-stimul
12	169.5	7.4	360	2 H70147	basic membrane pro
13	169	7.4	357	2 AD1248	CD4+ T cell-stimul
14	168.5	7.3	525	2 C82914	conserved hypotet
15	164	7.1	524	2 D82944	hypothetical membr
16	162	7.0	339	2 F70147	basic membrane pro
17	156.5	6.8	591	2 D64204	membrane lipoprote
18	147	6.4	353	2 G70147	basic membrane pro
19	143.5	6.2	384	2 T40867	hypothetical prote
20	141.5	6.2	539	2 D82886	conserved hypotet
21	141	6.1	379	2 H75318	membrane lipoprote
22	140.5	6.1	337	2 AH2591	membrane lipoprote
23	140.5	6.1	337	2 B97374	Deinococcus radiod
24	132	5.7	547	2 E29504	mercury(II) reduct
25	132	5.7	657	2 S73428	probable lipoprote
26	130	5.7	484	2 B90524	hypothetical prote
27	129.5	5.6	1140	2 S73786	hypothetical prote
28	129.5	5.6	1558	2 B71603	RESA-H3 antigen pr
29	127.5	5.5	763	2 A82863	hypothetical prote

30	127	5.5	349	2 F84246	hypothetical prote
31	126	5.5	326	2 G95857	hypothetical prote
32	125	5.4	1635	2 A10452	hemolysin (impor
33	124.5	5.4	556	2 H82301	peptide ABC transp
34	124.5	5.4	2269	2 T28677	thoxytry protein -
35	123.5	5.4	322	2 F84236	ABC transporter (l
36	123.5	5.4	626	2 C25035	colicin Ia - Esche
37	123.5	5.4	1031	2 C81302	probable type I si
38	123.5	5.4	1223	2 E88451	protein K10D2.1 (l
39	122.5	5.3	553	1 SMEBH1	flagellar hook-aaa
40	122.5	5.3	553	2 AH0640	flagellar hook-aaa
41	122.5	5.3	2285	2 T12796	probable transglyc
42	122.5	5.3	2346	2 T13829	tpi homolog - trui
43	121	5.3	1546	2 G90603	lipoprotein (impor
44	119.5	5.2	350	2 F70139	exported protein (
45	119.5	5.2	627	2 A41609	dnak-type molecule

ALIGNMENTS

RESULT 1

B90555
ABC transporter xylose-binding lipoprotein [imported] - Mycoplasma pulmonis (strain UAB
C:/Species: Mycoplasma pulmonis
C:/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:/Accession: B90555
R:/Chambaud, I.; Helling, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001.
A:/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:/Reference number: A99512; MUID:21267165; PMID:11353084
A:/Accession: B90555
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-461 <KOR>
A:/Cross-references: UNIPROT:Q96QL5; GB:AL445566; PID:G14089760; PIDN:CAC13519.1; GSPDB:G:
A:/Experimental source: strain UAB CTIP
C:/Genetics:
A:/Gene: MYPV 3460
A:/Genetic code: SGC3

Query Match	24.1%	Score 554.5	DB 2	Length 461
Best local similarity	30.3%	Pred. No. 2.6e-26		
Matches 140	Conservative 88	Mismatches 175	Indels 59	Gaps 13
QY	5	IKMKFGLGLVFPPLSAITISAGCMDEKTEKESKDNQNKOT--DVSKISGLVNER- 61		
DB	1	MKLNR--KLPSILFVALALALPATF-----VSCAQNPVKTNISLDSKITLDSQKE 51		
QY	62	-----KSEIMAKADANKHFGIANNATVYAGGTVDNSFNQSSWEAIOQALTGGEI 113		
DB	52	VTEYQKIVENKIKQASLEFQK-----VVLITADNDIDKSFNQVYSQKTLKQFVDKAY 106		
QY	114	TS-----VDSSTABIEGKYSLSIANTNNKVVWVLSGFQHDATFRWLKIDENK 159		
DB	107	KSQKKAENQHKLDNYINSVKDLQNYKVALDRGYTTLITGRQOENIEINFENDENL 166		
QY	160	QLPTEKNIITLIGDMWTEN-VIPTEGYINLTYTTERACMLAGANMSFLAKKPPSPPTK 218		
DB	167	RRFENKVKIIGVMAWPNANSKIPQGLISLIFTEBAGWQAGYASADFGTKYANNEAK 226		
QY	219	RSATVIGGIGSPATYDFIAGYLAGIKAMN--LKNSDKTKITTDKIEINLGFVDQSTK 276		
DB	227	RAIAPFGGDPAGTDTDLNGFEBIRKAMNSEAEKANKVKI VSNLVLDIGF-IPNKEK 285		
QY	277	ERLEQIASKDKPSTLLAVAGPLTEIFSDIT--ANQDRYLIGVTDQSLVYTKKPKF 333		
DB	286	EVANSNVETGKSTSLPVPAGPFTGVVDVLRKQTSDDRDRIIVGVDIQSLFTDSKRF 345		
QY	334	TSILKNGYGVFSVLSLDTY-----KKSNSNLAGFERGKSAVYVIGIKORFVDA 385		
DB	346	TSIVKNTIAPFYQTLALLTLTDEESVILKEGNDKFLDS---NPNKLVKKGSLAKFVNIT 402		

Qy 386 DTSLGNDKLTATEAISA--KKEFEKTKTIPAEVFKTIE 425
Db 403 KSRVKESTKQADSTICKAIDKMANPNSKKEEMTNGDLE 444

RESULT 2

G95097
110proteins [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: G95097
R/Reteljn, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: G95097
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-350 <KUR>
A/Cross-references: UNIPROT:Q97RH0; GB:AE005672; PIDN:AAK74976.1; PID:g14972319; GSPDB:G
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0845
C/Superfamily: ABC transporter ynfN

Query Match 9.7%; Score 223; DB 2; Length 350;
Best Local Similarity 24.8%; Pred. No. 2.4e-06;
Matches 110; Conservative 63; Mismatches 148; Indels 122; Gaps 22;

Qy 1 MKKKIKNNKFLGLGVPEPLSAIATIS-AGCWDKETTKEKSADNQKQITDVSKISGLVN 59
Db 1 MNKK----QWLGGLV----AVALVGLAACNRRSRNAASSDVYTK------ 39
Qy 60 ERKSEIMAAKADANKHFGLNMAIYTAGTVNDSFNQSWEAIOQLGA---LTGGEITSV 116
Db 40 -----AAIVDTGVDKDSFNQSAWEGLOAWGKEHNLSDKNGFTY 79
Qy 117 DSTAELE-----GKYSGLANTNNKVVWVLSGFQGDFAFTWMLKIPENKOLPTEKNI 167
Db 80 FQSTSEADYANNLQQAAGSYVLIIGV------GFLANNA-----YKDAKERTDILNY 125
Qy 168 IILGIDWTDENVLPTRGYINLTYKTEBAGWLAGYANASFLAKKPPSDPTKRSALVIGGG 227
Db 126 VLIDVDVTKDQKNV-----ASVTFADNESGYLAGVAAAK-----TTTKQGVFGG 170
Qy 228 I-SPAIVDFIAGYLAGIYANMLKNSDKTKITTDKIEINLGFVDQDTSKERLEQIASKD 286
Db 171 IESEVISRFEAGFRAGV-----ASVDPSTIKQVDVYAG-SFG-DAAKGTIAAQAAGAD 223
Qy 287 KPSTLLAAVAPL-TEIFSDIIANQND------YLIGVDTQSL-VYT-----KTKNKKF 333
Db 224 ---IVYQVAGGTGAGVFAEAKSLNESRPNKRWVIGDRQDEAGKTSKDGKESNFVL 280
Qy 334 TSILKNIGYVFSVLSDLTYTKKSNRLAGFEFGKSAATVYLGIDRFVADIADTSLGND 393
Db 281 VSTLKQVGTIVKDI-----SNKAERGEPGGQ---VIVYSKDKVDLAVTNLSSEG 329
Qy 394 KKLTAEASEAKKEFEKTKTIP 416
Db 330 KK-----AVEDAKAKILDSGVKVP 348

RESULT 3

C97965
conserved hypothetical protein spr0747 [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: C97965
R/Hokins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: C97965
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-374 <KUR>
A/Cross-references: UNIPROT:Q8DQC2; GB:AE007317; PIDN:AAK99551.1; PID:g15458340; GSPDB:G
C/Genetics:
A/Gene: spr0747

Query Match 9.5%; Score 219; DB 2; Length 374;
Best Local Similarity 24.6%; Pred. No. 4.7e-06;
Matches 109; Conservative 62; Mismatches 150; Indels 122; Gaps 21;

Qy 1 MKKKIKNNKFLGLGVPEPLSAIATIS-AGCWDKETTKEKSADNQKQITDVSKISGLVN 59
Db 25 MNKK----QWLGGLV----AVALVGLAACNRRSRNAASSDVYTK------ 63
Qy 60 ERKSEIMAAKADANKHFGLNMAIYTAGTVNDSFNQSWEAIOQLGA---LTGGEITSV 116
Db 64 -----AAIVDTGVDKDSFNQSAWEGLOAWGKEHNLSDKNGFTY 103
Qy 117 DSTAELE-----GKYSGLANTNNKVVWVLSGFQGDFAFTWMLKIPENKOLPTEKNI 167
Db 104 FQSTSEADYANNLQQAAGSYVLIIGV------GFLANNA-----YKDAKERTDILNY 149
Qy 168 IILGIDWTDENVLPTRGYINLTYKTEBAGWLAGYANASFLAKKPPSDPTKRSALVIGGG 227
Db 150 VLIDVDVTKDQKNV-----ASVTFADNESGYLAGVAAAK-----TTTKQGVFGG 194
Qy 228 I-SPAIVDFIAGYLAGIYANMLKNSDKTKITTDKIEINLGFVDQDTSKERLEQIASKD 286
Db 195 IESEVISRFEAGFRAGV-----ASVDPSTIKQVDVYAG-SFG-DAAKGTIAAQAAGAD 247
Qy 287 KPSTLLAAVAPL-TEIFSDIIANQND------YLIGVDTQSL-VYT-----KTKNKKF 333
Db 248 ---IVYQVAGGTGAGVFAEAKSLNESRPNKRWVIGDRQDEAGKTSKDGKESNFVL 304
Qy 334 TSILKNIGYVFSVLSDLTYTKKSNRLAGFEFGKSAATVYLGIDRFVADIADTSLGND 393
Db 305 VSTLKQVGTIVKDI-----SNKAERGEPGGQ---VIVYSKDKVDLAVTNLSSEG 353
Qy 394 KKLTAEASEAKKEFEKTKTIP 416
Db 354 KK-----AVEDAKAKILDSGVKVP 372

RESULT 4

F86804
basic membrane protein A [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86804
R/Bolotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: F86804
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-350 <STO>
A/Cross-references: UNIPROT:Q9CFM9; GB:AE005176; PID:g12724428; PIDN:AAK05536.1; GSPDB:G
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: bmpA

Query Match 9.0%; Score 207; DB 2; Length 350;
Best Local Similarity 24.2%; Pred. No. 2.3e-05;
Matches 107; Conservative 62; Mismatches 156; Indels 118; Gaps 20;

QY 1 MKKKIKNNKFLGLVPELPAIATIS---AGCWDKETTEKESADNONKOITDVSKI56
Db 1 MKKR-----VIAVSLALASVAVLAGCSHDAAGSGK-----32
QY 57 LVNERKSEIMAAKADANKHFGNLMAIYTAGTVDNDSFNOSWEAIOQLGALTG-----G 111
Db 33 -----AKTD-----LKAALIVTEIGVNDRFSNOSAMEGLQSGWKENNKKKGTG 75
QY 112 EITSVDSSTALEBKYSGLANTNNKVVVLISGFQHGDAFTRMLKIPENKQJTEBKNIILIG 171
Db 76 YTFPOSASDYTTNNVSAEQGYKLLFGTGSLSQDATTSAAK---NN---PKSNFYIVD 129
QY 172 IDWTDENVIPTRGYINLYTKTEBAGMLAGYANASFLAKKFPSPDKRSAL-VIGGGISP 230
Db 130 SVIQDQGV-----ASATFADNESAVLAGVAALK-----ATTKNKIGFICGMQSD 174
QY 231 AVTDPIAGYLAGIKAANLKNSSDKKTKITTDKIEINLGFVDVDTSTKERLEOIASKDPST 290
Db 175 VITFEKGYEKGAKSVN-----PDIKVYQVAGSFSDAKAGKTIAAAMYAGADVD 224
QY 291 LNAVAGPL-TEIFSDIIA-----NONDR-VLIGVDTQSLV--VT---KTKNKFPTSIL 337
Db 225 VYOCAGGVGTGVFEAKALNSTKNEADKVVIVGDQDEYKTKSKDKGDSNFWLVSTI 284
QY 338 KNLGYSVFVSLDLYTKKSNRNLAG--FEFGKKSATVYIGIKDRFYDIADTSLGNDKK 395
Db 285 KEVG---NVVKDADTKKOGKFGGRTIVYIDLNGGVNIGL-----DSANSEIK-----330
QY 396 LATEAISEAKKEFEKTKTTPAE 418
Db 331 ---DAVAKAKADIIDGKITVPSK 350

RESULT 5
D96986
Probable lipoprotein, Med/BMP family [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: D96986
R/Kolling, J.; Betton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:2159325
A:Accession: D96986
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <KUR>
A:Cross-references: UNIPROT:Q97L60; GB:AE001437; PIDN:AAK78679.1; PID:G15023581; GSPDB:C
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetic8
A:Gene: CAC0702

Query Match 8.5%; Score 195.5; DB 2; Length 357;
Best Local Similarity 22.3%; Pred. No. 0.00012;
Matches 100; Conservative 72; Mismatches 156; Indels 121; Gaps 20;

QY 1 MKKKIKNNKFLGLVPELPAIATISAGCWDKETTEKESADNONKOITDVSKI56
Db 2 IKKKT-----IAILTTVMIVAGIFAGC-----SSTSSGSGNS 33
QY 61 RKSEIMAAKADANKHFGNLMAIYTAGTVDNDSFNOSWEAIOQLGALTGGEITSVDSST 120
Db 34 K-----DTKK---VKVGLSTDEGLNDKSNFQAGADEQIKKAKEYSVIYRAIEBK 81
QY 121 A-ELEGKYSGLANTNNKVVVLISGFQHGDAF-TRMLKIPENKQJTEBKNIILIG 171
Db 82 KDDYQPNLQSLIDNDSDLVFGVGYQMDLDTIAKKYDPKFAIIDDAYKQPKNI-----137
QY 172 IDWTDENVIPTRGYINLYTKTEBAGMLAGYANASFLAKKFPSPDKRSAL-VIGGGISP 230
Db 138 -----MSLVFKEQSGSFLMG-----VLAGKM-----TKTNKIGFVGKQOP 173

QY 231 AVTDPIAGYLAGIKAANLKNSSDKKTKITTDKIEINLGFVDVDTST-KERLEQIASKDKPS 289
Db 174 LNKFLISGYIAGAKTVN-----PNITVEK---NTYNDSDISKKEVATSLYNGGCDI 223
QY 290 TLLAVAGPLTEISDIIANONDR-----YLIGVDTQSLVYTKKPKFTSILKNLGSVPF 345
Db 224 VYHAAGAGIGVF-DVAKSLRDQKDVMAIGVDDQAAGLPKPADVILTSMVKRVDIATY 282
QY 346 SVLSDLVTKKSNRNLAGFEFGKKSATVYIGIKDRFYDIADTSLGNDKLTATEAISEAK 405
Db 283 NTVKDLVKGK-----FEGKVES---FGLXEDGVGVAPTS---NKIVPSVSLVLD 328
QY 406 K-----EFEKTKTTPAEVR 421
Db 329 KYKKAIDGKIIVPDVDAKQTFKTDQIK 357

RESULT 6
C82946
Hypochemical protein U0012 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: C82946
R/Glass, J.I.; Letkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: C82946
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <GLA>
A:Cross-references: GB:AE002100; GB:AF222894; NID:G6898946; PIDN:AAF30417.1; GSPDB:GN001;
A:Experimental source: serovar 3, biovar 1
C:Genetics:
A:Gene: U0012
A:Genetic code: SGC3

Query Match 8.2%; Score 187.5; DB 2; Length 516;
Best Local Similarity 23.3%; Pred. No. 0.00061;
Matches 117; Conservative 67; Mismatches 167; Indels 151; Gaps 25;

QY 1 MKKKIKNNKFLGLVPELPAIATISAGCWDKETTEKESADNONKOITDVSKI56
Db 1 MKKLKKVFLFAGSVPALGTIITVATSCYQKSTL-----35
QY 61 RKSEIMAAKADANKHFGNLMAIYTAGTVDNDSFNOSWEAIOQLGALTGGEITSVDSST 120
Db 36 -----NYSQFYW-----TSPTSDD 49
QY 121 AE-LEGGYSLANTNNKVVVLISGFQHGDAFTRMLKIPENKQ--LFEK---NIIL---G 171
Db 50 DEGRQTKYSMASSGKALMLPGYH-----PERQNALVNDKFPBNLIALILDG 99
QY 172 IDWTDENVIPTRG--RYINLYTKTEBAGMLAGYANASFLAKK--PSPDKRSALVIGG 226
Db 100 VYNNDNKAEPYKGAADVAVFYKVDAAFLGIAAAVMLNSQAVFGAD---NKLTWGG 155
QY 227 --GISPA-VTDPIAGYLAGIKAANLKNSSDKKTK---ITTDKIEINL-----GPD 269
Db 156 YVGINAKNTNNYLAGFPLGVKMAEKKDKNIQOEGTETKRWIVNVOYVASSSSAGGFQ 215
QY 270 VQDSTKERLEQIASKDKPSLTLAVAGPLTEI-FSDIIANONDRY-LIGVD---TDQSL 323
Db 216 SDSANAKKIIOELITTK-ADLILPVALPOYGIANTEAIIATTSNHYGVIGVDEIENQAI 274
QY 324 VYTKNKKFPTSIL--KNLGYSVFVLS--DLYTKKSNRNLAGFEFGKKSATVYIGI- 377
Db 275 --NKTKDKFINTHLSGKN-GVIRFSITKRLDTTITLLENALIGESLSKSDIVIGSE 331
QY 378 ---KDRF---VDIADTSLG-----NDKLTATEAISEAKKEFEKTKTTPAE 418
Db 332 IDPKDKYKLGAVNTVGNLSDGVGISPSAHYVIDAFNLAAOTNOSDKVSTYDELVNKITND 391

QY 419 EVKRTLEIPENPKQPKQES 440
Db 392 DLFKTLDKKPYVDGYLDVKET 413

RESULT 7

ABC transporter (lipoprotein) homolog yuFn - Bacillus subtilis

C:Species: Bacillus subtilis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: C70009

R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Galle
lech, J.; Harwood, C.R.; Henaut, A.; Hiltbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.
koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Togonni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Dancin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: C70009

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-350 <KUN>

A:Cross-references: UNIPROT:005252; GB:Z99120; GB:AL009126; NID:G2635613; PIDN:CAB15143.

A:Experimental source: strain 168

C:Genetics:

A:Gene: yuFn

C:Superfamily: ABC transporter yuFn

Query Match 8.1%; Score 185.5; DB 1; Length 350;
Best Local Similarity 24.8%; Pred. No. 0.00046;
Matches 106; Conservative 61; Mismatches 149; Indels 111; Gaps 21;

QY 21 AIAITISACMDKETTERKSAADNOKQITDVSKISGLVNEKSEIMAKADANRFGIUM 80
Db 6 AAGTILGACGNSE--KSSGSGEGKXK-----RSV 32
QY 81 AIVTAGTIVNNSFNQSSWEAIOQLGALTG-----GEITSVDSSTAELEGYSIANTN 134
Db 33 AMVTDVGVDDKSPFQSGMEGIGARQKNGIKKNGNDYLOSQSDADYITNTAKLAEN 92
QY 135 KNWVLVSGFHGDAFTRLKIPENKQIFTEKNIIILGID-WTDTENVIPTGRYINLTYYT 193
Db 93 FDLIVGVGLMEDSIS--EIAIDQR-----KNTNPAIIDAVIDKONV-----ASITPE 138
QY 194 EAGGLAAYANASFLAKKPPSDPTKRSAL-VIGGISPAVVDPIAGIAGIKAMLNKNSD 252
Db 139 QEGSLVVAAL-----SSKSGKIGFVGMESELKKFEVGFAGVAVP-K-- 186
QY 253 KKTITTKIKRINILGFVDQT--STKERLEQIASDKPSTLLAVAPL-TEIFSDIAN- 308
Db 187 ---AVEVYKAGFDKADVGAFAESM---YKSGDVLYHSAGAGTGVTFE-AKUL 236
QY 309 ---QNDRYILGVDTQD--SLVYTKTKNKFETSIKLNIGSVSVSLSDLYTKKSNRN 360
Db 237 KKEDPKRDVWVIGVQDKQYAEQVGEQTDNVTLTSMVK-----KVDYVEDV-TRKASDGK 291
QY 361 LAGPFGKSAIVYIGIDRPVIDADTSLBGNDKLALEAISAKKEP-----FEK 411
Db 292 PPGGS-----TLTYGLDQDVGISPSKONLSDDVI--KAVDKMKKIIDGLEIPATEKE 343
QY 412 TKTIPAE 418
Db 344 LKTFKAE 350

RESULT 8

H71340

membrane lipoprotein Tmpc precursor - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 07-Aug-1998 #sequence_revision 07-Aug-1998 #text_change 09-Jul-2004

C:Accession: H71340; A43595; S29561

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ueteback, T.; McDo
rthy, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:9832770; PMID:9655876

A:Accession: H71340

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-353 <COL>

A:Cross-references: UNIPROT:P29724; GB:AE001211; GB:AE000520; NID:G3322582; PIDN:AAC6530;

A:Experimental source: Strain Nichols

R:Schouls, L.M.; van der Heide, H.G.J.; van Embden, J.D.A.

Infect. Immun. 59, 3536-3546, 1991

A:Title: Characterization of the 35-kilodalton Treponema pallidum subsp. pallidum recomb;

A:Reference number: A43595; MUID:91372962; PMID:1894360

A:Accession: A43595

A:Molecule type: DNA

A:Residues: 1-10, 'A', 12-158, 'R', 160-353 <SCH>

A:Cross-references: GB:X57836; NID:G48838; PIDN:CAA40968.1; PID:G581809

A:Note: this protein is shown to incorporate palmitic acid

C:Genetics:

A:Gene: tmpc; TP0319

A:Start codon: GTG

C:Superfamily: ABC transporter yuFn

C:Keywords: blocked amino end; lipoprotein; membrane protein; thiolester bond

F.1-20/Domain: signal sequence #status predicted <SIG>

F.21-353/Product: membrane lipoprotein tmpc #status predicted <MAT>

F.21/Modified site: fatty acylated amino end (Cys) (in mature form) #status predicted

F.21/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status predicted

Query Match 7.9%; Score 182; DB 1; Length 353;
Best Local Similarity 23.7%; Pred. No. 0.00077;
Matches 86; Conservative 69; Mismatches 134; Indels 74; Gaps 17;

QY 80 MAITTAGTIVNNSFNQSSWEAIOQLGALTGGEITSVDSST-ALEBKYSSLANTNKVM 138
Db 42 VGMVTDSDGIDDKSFFNOQVWEGISRFQENNAKCKYTAIDAYVPSLSAFADENNGLV 101
QY 139 VLSGFQHGDAFTRLKIPENKQIFTEKNIIILGIDWTDTENVIPTGRYINLTYYEAGW 198
Db 102 VACGSFLVEA-----VIFTSARFPKQFLVIDAVVDNRDV-----VSAVFGQNGSF 149
QY 199 LAGYANASFLAKKPPSDPTKRSAL-VIGG--GISPAVTDPIAGIAGIKAMLNKNSDK 254
Db 150 LVGVA--AALKKE-----AGKSAVGPIVGMELGMPPL--FEAGFEAGVAVD----- 193
QY 255 TKITTDKIEINLGFVDQT--STKERLEQIASK--DKPSTLLAVG-----PLTIFSDII 306
Db 194 -----PDIOVVEVANTFSDPQGOALAALYDSGNVYFOVAGGTGNGVIREADR 246
QY 307 ANQNDRLIAGVDTQDQSH--VYTKTKNKFETSIKLNIGSVSVSLSDLYTKKSNRNLAGF 364
Db 247 LMGDVWVIGDRKQYMDGYDSKSVLTSWVRA-----DVAABKISKMAVYGSFPG- 300
QY 365 EFGKKSATVYIGIDRPVIDADTSLBGNDKLALEAISAKKEPTEKTKTIPAEVKTLL 424
Db 301 ---GGSIMFGLDKAVGIPF-----ENPNLSSAVMEKIRSPREKI-----VSKEI 342
QY 425 EIP 427
Db 343 VVP 345

RESULT 9

F72418

EMBL Microbiol Lett. 119, 381-388, 1994
A:Title: Nucleotide sequence and analysis of the gene in *Borrelia burgdorferi* encoding the
A:Reference number: 140289; MUID:94327086; PMID:8050720
A:Accession: J140290
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: J140290
A:Molecule type: DNA
A:Residues: 1-232, 'A', 234-317, 'V', 319-341 <RES>
A:Cross-references: GB:J14194; NID:9508420; PIDN:AAA72407.1; PID:g508422
R:Jain, C., Davidson, B.E., Saint Girons, I., Old, I.G.
Microbiology 140, 2931-2940, 1994
A:Title: Conservation of gene arrangement and an unusual organisation of rRNA genes in
A:Reference number: 140241; MUID:95111614; PMID:7812434
A:Accession: J140242
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-44, 'A', 46-179 <RES>
A:Cross-references: GB:J15050; NID:9516591; PIDN:AAC41402.1; PID:g551744
C:Superfamily: basic membrane protein C

Query Match 7.7%; Score 178; DB 2; Length 341;
Best Local Similarity 26.7%; Pred. No. 0.0013;
Matches 100; Conservative 60; Mismatches 123; Indels 92; Gaps 25;

QY 72 ANKRGHLMVAITACGTVVNDNSFNQSSWEAIOQLGALTGETSVDSSTAELEKYS-- 129
DB 24 SSKKIKIMLV--DGVLDKSPNSANELLRLKKDPENIEVFS--CAISGVSSYV 78
QY 130 --LANTNKN---VWVLSGFQHGDAFTRWLKIPEKOLFTEKNII--ILGIDWTDTE NVI 181
DB 79 SDDLRLKNGSDLIW-LVGYMLTDA--SLVSESEPKI--SYGIDPIYGDVQIPEWLI 133
QY 182 PTGRINLTYTEEAGMLAGY--AANSPFLAKKRPSPDKSAIVIG--GGSPAVTD-FI 236
DB 134 A-----VVFREQAQFLAIGYIAAKKSFSGK-----IGRIGGKGNIVADF 174
QY 237 AGYLAGIVAMVKNKSDDKTKITTDKIEINLGFVQDTSTKERLEQIAKSK--DKPSTLLAV 294
DB 175 YGESGAKYAN-KDIEIISSEYNSFSVDVIG-----RTLSKMYSKGIDVHF 221
QY 295 AGPLTEI-FSDIIANQND-RYLIGVDTDOSLVYTKTKNKEFTSLKNGSVFSVLSLY 352
DB 222 AAGLAGIVIEPRAKNLGDSYVYIGADQDS--YLAPEKN-FITSYIKNIGDALYLTGE-Y 277
QY 353 TKKSNRNLAGPEFEGKSKATVYLGIDRFVDIADTSLSEGNDKKLATPAISEAKKEPEKT 412
DB 278 IKNNNV-----WEGGK--VVGQGLRDGVIGLPAN-----EF-EYI 310
QY 413 KTIPEAEVRKLTLEIP 427
DB 311 KVLERRIKNKETIVP 325

RESULT 11
AH1610
CD4+ T cell-stimulating antigen, lipoprotein [imported] - *Listeria innocua* (strain C1p1
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1610
R:Glaser, P.; Frangenul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entlian, K.D.; Fsthi, H.
Science 294, 849-852, 2001
A:Authors: Kretz, U.; Kuhn, M.; Kunat, F.; Kuzepkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Smeets, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <GLA>
A:Cross-references: UNIPROT:Q928W7; GB:AL592022; PIDN:CAQ96656.1; PID:g1641389; GSPDB:GN
C:Experimental source: strain C1p11262
C:Genetics:

A:Gene: tcea
C:Superfamily: ABC transporter ynfN

Query Match 7.6%; Score 174; DB 2; Length 357;
Best Local Similarity 24.2%; Pred. No. 0.0024;
Matches 101; Conservative 62; Mismatches 137; Indels 118; Gaps 23;

Qy 55 SGLV-----NKKSEIMAKADANKHFGALMAIVTAGTVNDNSFNQSSWEALIQGL 106
Db 16 SGLV-----NKKSEIMAKADANKHFGALMAIVTAGTVNDNSFNQSSWEALIQGL 106
Qy 107 ALT-----GGLTSDSSTAELEGKYSILANTNKVWLS-----GFGDGAFTRLK 154
Db 70 KANMEKGTIDGYNIQSSEADYK-----TNLNTAVRSVDLYIGIGYKLDALIEVSK 123
Qy 155 -IPENKLFTEKNIIILIGIDWTDTEENVIPTRGYINLYKTEBAGWLAGYANASFLAKFP 213
Db 124 QKPNQFPIVDDTL-----DNRNVYSIG-----FKMDGSLYGVAGL----- 163
Qy 214 SDPTKSAIVIGGISPAVTD-FLAGLAGIKAMNLKNSDKTKITTDKIEINIGFDVQ- 271
Db 164 --TTKTNKVGFGVGKAVIDRFEGAFYAGVKA-----VNPNAQIDVQY 205
Qy 272 --DPTKRLKQIASKDKPS-----TLAVAGPLTEIFSDITANQNR-----YLIGVDT 319
Db 206 ANDPAKADKGOIASSMYSSGVVDFHAGGTGNGVFAL-AKNLKKQPSRAVWVIGYDR 264
Qy 320 DQ-----SLVYTKNKFPTSLIKNLGYSVFVSLDYTKKNSRNLAG-FEFGKKSAT 372
Db 265 DQMBGKATANDGADYNTLTSEIKRVDIIV-----DLAT-----RTAAGFPFGTK--- 312
Qy 373 YLIGIKRFDIADTSLGNDKLTATEAISAKKEFEKTKTTPAEVRYKLTLEIPMP 430
Db 313 IEYGLDKAVGLSE-----HODNISKDVLAKE--EYKQKIVDGD-----IKVPEKP 357

RESULT 12

H70147
basic membrane protein D (bmpd) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C:Accession: H70147
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lachisra, R.; White
son, D.; Peterson, J.; Kervilavag, A.R.; Quackenbush, J.; Salzbey, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70147
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-360 <KLE>
A:Cross-references: GB:AE001144; GB:AE000783; NID:g2688291; PIDN:AA891505.1; PID:g268829
A:Experimental source: Strain B31
C:Superfamily: basic membrane protein C

Query Match 7.4%; Score 169.5; DB 2; Length 360;
Best Local Similarity 25.4%; Pred. No. 0.0045;
Matches 88; Conservative 64; Mismatches 137; Indels 57; Gaps 20;

Qy 65 IMAAKADANKHFGALMAIVTAGTVNDNSFNQSSWEALIQGLTGEITSVDSSTAELE 124
Db 34 VACSSSDGKSEAKTVSLI-VDGAFDDKGFNESSKAIKRLKADLNINIEKASTGNSYL 92
Qy 125 GKSSSLANTNKG-VWVLSGFQHGDAFTMLKIPENKQLFTEKNIIILIGIDWTDTEENVIP 183
Db 93 GDINLIEGNSNLWGI-GFRLSDIL--FORASEVAVS---NVAI--IEGVYEIOIPLK 143
Qy 184 GRVYNLYKTEBAGWLAGYANASFLAKKFPSPDPTKSAIVIGGISPAVTD-FLAGYLAG 242
Db 144 -NLNLSIRSEBVAFLAGY-----FASKASKTGKIGFVGVGKAVLSEFMIGYBAG 193

Qy 243 IKAMNLKNSDKTKITTDKI-----EINLGFVDVDTSTKERLEQIASDKPSTLLAVAGPL 298
Db 194 AKYAN-----SNIKVYSQYVTFPFGFGLG-----RSTASNM---YBDGVDIIFPAG-L 238
Qy 299 TELFSDIILANO-NDRVLIGVDDQSLVYRTKKNKFTSLIKNLGYSVFVSLDYLTRKS 356
Db 239 SGIGVITAEAKELGPDHYIIGVDQDQSYL---APNNVYSAVKKVDLSLMSL-----TKKY 290
Qy 357 NSRNIAFEFGKKSATVYLGK-DRFVDIADTSLGNDKLTATEAI 401
Db 291 LETGV--LDGSK---TWFLGKEDGLVLNENLKNYSSEIYNKSL 331

RESULT 13

AD1248
CD4+ T cell-stimulating antigen, lipoprotein [imported] - Listeria monocytogenes (strain
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1248
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entlian, K.D.; Feihl, H.,
Science 294, 849-852, 2001
A:Authors: Kreft, U.; Kuhn, M.; Kunz, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A>Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1248
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <GLA>
A:Cross-references: UNIPROT:Q48754; GB:NC_003210; PIDN:CAC9466.1; PID:G16410817; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: tcea
C:Superfamily: ABC transporter ynfN

Query Match 7.4%; Score 169; DB 2; Length 357;
Best Local Similarity 23.8%; Pred. No. 0.0048;
Matches 98; Conservative 63; Mismatches 144; Indels 106; Gaps 22;

Qy 54 ISGLVNRKSEIMAKADANKHFGALMAIVTAGTVNDNSFNQSSWEALIQGLT----- 109
Db 19 ILGACSSSDDKSSDDKSKDF--TVAMVTDGVDNRSFNQSSAWGGLKFGKANDMEK 76
Qy 110 GGLTSDSSTAELEGKYSILANTNKVWLS-----GFGDGAFTRLK-IPENKQ 160
Db 77 GTDGYNTLQASSEADYK-----TNLNTAVRSVDLYIGIGYKLDALIEVSKQKPNQF 130
Qy 161 LPTKKNIIILIGIDWTDTEENVIPTRGYINLYKTEBAGWLAGYANASFLAKKFPSPDKRS 220
Db 131 AIVDDTL-----DNRNVYSIG-----FKMDGSLYGVAGL-----TTKTN 168
Qy 221 AIVGGISPAYVD-FLAGLAGIKAMNLKNSDKTKITTDKIEINIGFDVQ--DPTK 276
Db 169 KVGFGVGKATVIDRFEGAFYAGVKA-----VNPNAQIDVQYANDPAKA 212
Qy 277 ERLKQIASKDKPS-----TLAVAGPLTEIFSDITANQNR-----YLIGVDTQ----- 321
Db 213 DKQGLASSMYSSGVVDFHAGGTGNGVFAL-AKNLKKQPSRAVWVIGYDRQMBEGK 271
Qy 322 -SLVYTKNKFPTSLIKNLGYSVFVSLDYTKKNSRNLAG-FEFGKKSATVYLGK 379
Db 272 VTANDGADYNTLTSEIKRVDIIV-----BDLAT-----RAKAGDFPFGTK---IEYGLDK 319
Qy 380 RFDVIADTSLGNDKLTATEAISAKKEFEKTKTTPAEVRYKLTLEIPMP 430
Db 320 DAVGLSE-----HODNISKDVLAKE--EYKQKIVDGD-----IKVPEKP 357

RESULT 14

C82914
conserved hypothetical membrane lipoprotein U0226 [imported] - Ureaplasma urealyticum

C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: C82914
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
 Submitted to GenBank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit
 A/Reference number: A82870
 A/Accession: C82914
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1525 <GUA>
 A/Cross-references: GB:AE00123; GB:AF222894; NID:G6899223; PIDN:AAF30635.1; GSPDB:GN001
 A/Experimental source: serovar 3; biovar 1
 C/Genetics:
 A/Gene: UU226
 A/Genetic code: SGC3

Query Match 7.3%; Score 168.5; DB 2; Length 525;
 Best Local Similarity 21.0%; Pred. No. 0.0089;
 Matches 113; Conservative 77; Mismatches 168; Indels 181; Gaps 24;

```

QY 3 KKIWMKFLGIGLVPPLSAITATISAGCWDKTTKEKSADNONQOITDVSKISGLVNERK 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 RKILDKKIIILASVIGIMGVALVPAL-----VACSONTKSDSRMLT----- 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 SEIMAAKADANKHFGMLMAIVTAGTVNDNSFNOSWEALIOQLALTGGEITSDSSYAE 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 53 -----LNPFQ-----RPSDGDGAQAI----- 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 LEGRYSSSLANTNNKVVWLSGFQGDFAFTRMLKIPENKOLFTEKNIIILGIDWTDTEV 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 72 -----YASKIODGAMLGILISFRHKNPISIKYFNSPKDNQ---QVSAVLI---DEIYDLQ 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 TG--RYINLTYYKTEEAGMLAGYANASFLAKK---FPDPTKRSALVIGGGIS---PAVTD 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 120 TGDRIASITVRADQAFPLGIAAAYLANSQNVFGKD---NKLTWGGFVGLHLPSTTR 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 235 FIAGYLAKIKAMNLKNSDKTKITTDKIE---INL-----GPDVQDSTYKER---LE 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 FIOGFKGCIOMAKLKNKKYKOTENNEKEKWINEVOFAFVYOSGDFSPISDAKAIYN 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 QIASKDPSTLLAVAGPLTEIFSDIIANQND--YLIGVTDQSLVYTKTKNKKF--TSIL 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 QLVAS--NNVDLILPVAGPDIVATTAABSSKPIVVGVDTEQELDDMTNKRARISENNKSL 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 KNLGYSVFSVL--DYTK-----KSNRNLAGFEFGKKSATVYIGIDRFYDI 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 295 ANGKTIIFSIYKRLDLAFKGLALKASGSAQLTNDINKDAYKLGHTTEASF--NKATYVD- 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 385 ADTSL-----EGNDKCLATEAISE-----AKKEPEEKT 412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 352 -NTALVELSKAGHOYLDAIKLSGLKEVNDKTIIVEIIQEDPLFKLSQIQTKLDEYAT 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 KTIIPAEVRKTLLEIPMP-----DKQDPKQOES 440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 KSOQGDWVLKS--EYQDLPFIQLQKMLGGLVYVDQKNELYPELNSFYLEKDPNKRQAS 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

D82944
 hypothetical membrane lipoprotein UU016 [imported] - Ureaplasma urealyticum
 C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: D82944
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
 Submitted to GenBank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit
 A/Reference number: A82870
 A/Accession: D82944
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-524 <GUA>
 A/Cross-references: GB:AE00101; GB:AF222894; NID:G6898957; PIDN:AAF30421.1; GSPDB:GN001

A/Experimental source: serovar 3; biovar 1
 C/Genetics:
 A/Gene: UU016
 A/Genetic code: SGC3

Query Match 7.1%; Score 164; DB 2; Length 524;
 Best Local Similarity 21.0%; Pred. No. 0.017;
 Matches 108; Conservative 73; Mismatches 178; Indels 156; Gaps 21;

```

QY 3 KKIWMK---FLGIGLVPPLSAITATISAGCWDKTTKEKSADNONQOITDVSKISGLV 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 KKSINKKIILFASLAGVVTILSVAAIASCND-----SNDDGKTNDGNYISKLS----- 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 ERKSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNOSWEALIOQLALTGGEITSDSS 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 53 ---LEDFYAKPAGDSDSGVHRT----- 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 120 TAELEGRYSSSLANTNNKVVWLSGFQGDFAFTRMLKIPENKOLFTEKNIIILGIDWTDTE 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 72 -----YNSLYDDGAMLGILISFHSIPIKEYFGSSSDKK---DLGAVLIDDKFSGT-- 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 VIPTGRYINLTYYKTEEAGMLAGYANASFL---AKKPSDPTKRSALVIGGGIS---PAV 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 120 -VGKRIASVSYRVDQAFPLGIAAAYLANSQNTFPAAD---GKLTWGGVGLHFTST 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 DFIALYLAGIKAMNLKNSDKTKITTDKIEINLGFVDQDSTYKER---LEQI-ASK--- 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 175 TFIQGFPLGQVMNEKIKDK-----EIN-----QEDANGSKKMMANVEQVASKYVA 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 -----DKPSTLLAVAGPLTEIFSDIIANQND--RYLIGVTDQSL-- 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 222 GSFKPDDEGATNIINDLITTKADVILPVAGPOTMLATSIYSNATDPSEVIIIGVDTAOEILD 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 324 -----VYTKQNKFF---SILKNLGYSVFSVSLDY--TKSNSNLAGFEFGKKSAT 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 282 VTNKRRTTNKTVNDGKTIPLFSIVRVDLAMKALENSKAQQLTNDINKDAYKLGHTTEA 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 VYLGIDRFYDIADTS-----LEGNDKCLATEAISEAK-----KEFEE 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 342 SL-----DKSTYVDVDTPLVELSNAGRVYLEQAALAGAKATITVAQIVVAVIQNEELFKL 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 KTKTIIPAEVRKTLLEIPMPDKQDPKQOESLDKLI 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 398 KGTFLKEDLATKTSIDGWVLKDSSEKNSFSELOKLL 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: December 18, 2004, 01:27:29
 Job time : 30.349 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:14:44 ; Search time 127.457 Seconds
(without alignments)
2035.941 Million cell updates/sec

Title: US-09-676-249D-2
Perfect score: 2299
Sequence: 1 MKKKIKMKFKLGLVFLPLS.....KPPDKQOESLKLITDINKN 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	706	30.7	457	2	Q48902 mycoplasma
2	554.5	24.1	461	2	Q980L5 mycoplasma
3	467.5	20.3	491	2	Q6K104 mycoplasma
4	467.5	20.3	491	2	AAT27522 mycoplasma
5	438.5	19.1	468	2	Q6E5B5 mycoplasma
6	417	18.1	465	2	Q9X775 mycoplasma
7	400.5	17.4	428	2	O52311 mycoplasma
8	397.5	17.3	428	2	Q9RGX5 mycoplasma
9	394.5	17.2	428	2	O32417 mycoplasma
10	394	17.1	428	2	Q9RGX4 mycoplasma
11	394	17.1	428	2	Q9RGX6 mycoplasma
12	393.5	17.1	428	2	Q9RGX7 mycoplasma
13	389.5	16.9	428	2	Q9RGX3 mycoplasma
14	387.5	16.9	428	2	Q9RGX6 mycoplasma
15	247	10.7	352	2	Q6RCY2 mycoplasma
16	238	10.4	457	2	Q6F0E8 mycoplasma
17	232.5	10.1	350	2	Q992H4 mycoplasma
18	229.5	10.0	349	2	Q9EV91 mycoplasma
19	227.5	9.9	349	2	O878P8 streptococc
20	226.5	9.9	349	2	O8DUJ6 streptococc
21	226.5	9.9	350	2	O8K7C9 streptococc
22	224.5	9.8	350	2	O8P0W2 streptococc
23	223	9.7	350	2	Q97RH0 streptococc
24	219	9.5	374	2	Q8DQC2 streptococc
25	215.5	9.4	349	2	O8DZY8 streptococc
26	215.5	9.4	349	2	O8ESN9 streptococc
27	211	9.2	359	2	O898S3 clostridium
28	210	9.1	361	2	O81A05 bacillus ce
29	207	9.0	350	2	Q9CFM9 streptococc
30	197	8.6	347	2	O8E1F0 streptococc
31	197	8.6	347	2	O8E6W8 streptococc

32	197	8.6	466	2	Q7WU11 spiroplasma
33	195.5	8.5	357	2	Q97L60 clostridium
34	194.5	8.5	550	2	Q6MUL7 mycoplasma
35	194.5	8.5	550	2	CAE76664 mycoplasma
36	191.5	8.3	347	2	Q99YMS streptococc
37	191.5	8.3	347	2	O8K8S6 streptococc
38	190.5	8.3	357	2	Q83919 enterococcu
39	188	8.2	368	2	O8FXQ5 bruceella su
40	187.5	8.2	347	2	O8P2U0 streptococc
41	187.5	8.2	516	2	Q9PRD3 ureaplasma
42	185.5	8.1	350	1	YUEN_BACSU bacillus ce
43	185.5	8.1	355	2	Q732T4 bacillus su
44	185.5	8.1	355	2	AAS42731 bacillus
45	184	8.0	325	1	BMPA_BORGA borrelia ga

ALIGNMENTS

RESULT 1

ID	Q48902	PRELIMINARY;	PRT;	457 AA.
AC	Q48902;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Ag 243-5 protein precursor.			
OS	Mycoplasma arginini.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_TaxID=2094;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96163149; PubMed=8551970;			
RA	Uehio S., Iwaki K., Tanai M., Ohta T., Fukuda S., Sugimura K.,			
RA	Kurimoto M.,			
RT	"Metabolic promoting activity of a novel molecule, Ag 243-5 derived from Mycoplasma and the determination of the complete nucleotide sequence."			
RT	Microbiol. Immunol. 39:393-400(1995).			
RL	EMBL; D16674; BAA04082.1; -			
DR	GO; GO:0008289; F:lipid binding; IEA.			
DR	InterPro; IPR003760; Bmp.			
DR	Pfam; PF02608; Bmp; 1.			
DR	PRINTS; PR01733; LIIPROTEIN48.			
KM	Signal.			
FT	SIGNAL 1 25 Potential.			
FT	CHAIN 26 457 Ag 243-5 protein.			
SQ	SEQUENCE 457 AA; 50789 MW; 9AEEBA11620CB22F CRC64;			
Query Match 30.7%; Score 706; DB 2; Length 457;				
Best Local Similarity 38.4%; Pred. No. 2.7e-35;				
Matches 173; Conservative 87; Mismatches 145; Indels 46; Gaps 16;				
QY	14 GLVFLPSAIAIT-ISAGCWDKETTKERKADNOKQITDVSKISGLVNERKSEIWAAYADA 72			
DB	12 GTISTVASVATFVSCG---ETDKKGLI-----RIDP-----NSFVDRDAIEKA----- 54			
QY	73 NKHFGLMAIVTAGTVNDNSFNQSSWEAI-----QQLGALTG-GEITSVDSSTAELEGY 127			
DB	55 -KNDFEFTVLITAGGTVDKSPFNQSIWEAVLEHYDQIEKTTNLDNRVQETNNOSELGKY 113			
QY	128 SLSAANTNKNVWVLSGPHGDAFTRMKIPEN-----KQLTEKNIIILIGDW---TGTEN 179			
DB	114 KNFLNGKNVWVILTFGQGEFFPEFLKQTSNGKRYSDLAIEKVVITVAVDMDLSKEDKD 173			
QY	180 VIPTGRYINTLYTKEAGWLAGYANASFLAKKPPSPDPTKSAIVTGGISPAVDTFIAGY 239			
DB	174 LIRAGHRTISILYKTEBAGFIAGYASSKFLAYKFPNDAKRTIAPFGGHHGAGVDTFIAGF 233			
QY	240 LAGIKANLKNKSDKRTKITTDKIEINIGPDVDTSTYKERLEQJASKDKPSTLAAVAPLT 299			
DB	234 LAGIAKYNNDNPTAKVTISDNNINIDTGF-ISNDKTATFTINGIVNKS--SLIVFVAGSLT 290			

QY 300 EIFSDII--ANQNDRIYLGVDTPDQSLVYTKRKNKFTSILKNLGYSVFSVLSIDLYTKSN 357
D 291 SSVVDALIKSNKDKRYLIGVDTPDQSKI FSPA-TWFFSTIEKHLSRTIYQVLTIDWLKED 349
QY 358 SRNLAGEFGK----KSAIVYLGIKDRFVUDADNLSLENDKKALTEALSEAKKEFEETK 413
D 350 SKFLGSPFSFLTNPAVATYKGISDDFVGVNSTVADADYKVAQOEFLNEATDFKQIQ 409
QY 414 TIPAEVRKLTLEPEM--PDKOPDKQESL 441
D 410 ANPT-NYKSVLGIPITMLINDADAKNEKSL 439

RESULT 2
Q980L5 PRELIMINARY; PRT; 461 AA.
ID 0980L5
AC 0980L5
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE ABC TRANSPORTER XYLLOSE-BINDING LIPOPROTEIN.
GN OrderedLocustNames=MYPU_3460;
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP.
RX MEDLINE=21267165; PubMed=11353084;
RA Chazambad I., Heilig K., Fietris S., Barbe V., Samson D., Gallsen P.,
Mozer I., Dybvig R., Wroblewski H., Viart A., Rocha E.P.C.,
Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
Mycoplasma pulmonis."
RL Nucleic Acids Res. 29;2145-2153 (2001).
DR EMBL; AL445564; CAC13519.1; -.
DR PIR; B90555; B90555.
DR MYDULIST; MYPU_3460; -.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR InterPro; IPR008107; Mycoplasma_p48.
DR InterPro; IPR011050; Pectin_lyas_like.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 461 AA; 51096 MW; 6724D4D820809CE4 CRC64;

Query Match 24.1%; Score 554.5; DB 2; Length 461;
Best Local Similarity 30.3%; Pred. No. 6e-26;
Matches 140; Conservative 88; Mismatches 175; Indels 59; Gaps 13;
QY 5 IKMKKFLGLVPLSLAATISAGCWDKETKEKSDANONKQIT--DVSKISGLVNER- 61
D 1 MKLNK--KLRSILFLVAALALALPATF-----VSCAQNPVKTNNSLDSSKITDLSQKE 51
QY 62 -----KSRIMAKADANHFGLNMAIVTAGGTVDNNSFNQSSWEALIQOLGALTGEI 113
D 52 VTEQKIVENKIKQASLETQK-----VVLITADNIDKSFNQGYBQKTLKQFVDKAY 106
QY 114 TS-----VDSSTAELBQKYSLSANTNKNVWVLSGFQHGDAFTRMILKIPENK 159
D 107 KSQKKEAKENQKLDNINSAVKDLEQNYKVALDRGYTTWILTFQOGGEIENFINDENNL 166
QY 160 QLFTEKNIIILGIMWTDEN-VIPGRYINLTYTEEAGWLAGYANASFLAKKEPSDPTK 218
D 167 RRFENKKNKIIGVMAVPAVANSKIPOGSLISLFTTEEGWQAGYASADFLGTKANNEAK 226
QY 219 RSAIVIGGIGSPATYDFIAGYIAGIKAWN--LKNSDKTKKITTDKIEINLGFVDQDSTK 276
D 227 RAISAFGGGDFAGYTDPLNGFFBGRAMNSAEANANKVKYVSENVLVDITGF-IPNAEK 285
QY 277 ERLEQIASKDKRPSLTLLAVAGPLTEIFSDI--ANQNDRIYLGVDTPDQSLVYTKRKNK 333

D 286 EVSNVAVETGKSTISLVAAGFTGVVDVLRKDTSDERFIQVDTQSLVTSIDSKRFP 345
QY 334 TSILKNLGYSVFSVLSIDLYT-----KKSNSRLNLAFFEGKKSATVYLGIKDRFPVIA 385
D 346 TSIYKNIAFPVYQQLALLRLKDESVYLKKGNDKFLGS---NPNKVLKKGSLAKFPVNT 402
QY 386 DTSLEGNDDKALTEALISEA--KKEFEETKTIIPAEVRKLTLE 425
D 403 KSRVKEISIKTOADTSTIQKAIKMANPNSKKIEKEMTNCBLE 444

RESULT 3
Q6KIQ4 PRELIMINARY; PRT; 491 AA.
ID 06KIQ4
AC 06KIQ4
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE Unspecified sugar ABC transporter binding protein.
GN OrderedLocustNames=WMOB0360;
OS Mycoplasma mobile.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=163K / ATCC 43663;
RA Birren B.W., Strange-Thomann N., Smith C., Decaprio D., Fisher S.,
Butler J., Calvo S., Elkin T., Fitzgerald M.G., Haefer N., Kodira C.D.,
Major J., Wang S., Wilkinson J., Nicol R., Nusbbaum C.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR InterPro; IPR008107; Mycoplasma_p48.
DR PRINTS; PR01733; LIPOPROTEIN48.
KW Complete proteome.
SQ SEQUENCE 491 AA; 52817 MW; 2D1C707687771B41 CRC64;

Query Match 20.3%; Score 467.5; DB 2; Length 491;
Best Local Similarity 29.7%; Pred. No. 1.5e-20;
Matches 141; Conservative 80; Mismatches 165; Indels 89; Gaps 18;
QY 43 NQNKQITDVSKISGLVNERKSEIIMAKADANKRGLNMAIVTAGGTVDNNSFNQSSWEAL 102
D 32 NTGRLRVTDNQVFRDLVASREAFATQRVANNLSFNSTLTLLITLGGVNDLSPNQSINEAL 91
QY 103 QOLGALTG--GEITSVVSSTA---ELEBKYSLSANTNKNVWVLSGFQHGDAFTRMILKIP 157
D 92 LEIRQTKKGNPSPFATTTAGTPOQLORYDQALFFNHKFWVLTFQODGAFQWLDIGN 151
QY 158 NKQLFTEKNIIILGIMWTDENVIPGRYINLTYTEEAGWLAGYANASFLAKKEPSDPT 217
D 152 NRAEFIRKQYIIVADWTNLEIVPPQGFISINVRTOSSWIVGNAAKFIIDHNHNN-- 209
QY 218 KRSIVIGGIGSPATYDFIAGYIAGIKAWNLS-----DKTKKITP-DKIEI 264
D 210 -RTFNTFGGAFPPVIVNFNAGFLQGLIDFN--NSTFLPEBETSTITDKKUSFTPGDINI 266
QY 265 NLGFVDQDSTKERLEQIASKDKRST--LLAVAGPLTEIFSDIAND--RYLIGVDTQ 321
D 267 NTGPAV---TPBAATLQISIVSGTQVFPVAGSLTTLTVNLSIQNSQGFVIGVDSQ 322
QY 322 SLVYTKRKN--FTSILKNLGYSVFSVLSIDLY-----TKKNS 358
D 323 AKAFSPDLALTLFSSVEKNVAGTYAALASLYIGTAVTDPFNFITGSSRFIPYTERGNS 382
QY 359 RNLAGFEFGKKSATVYIGIDR-----FVDIADTSLBGNDDKALTEA-----ISEAKEF 408
D 383 SSIP-----LANADITGFVFESTEPVDVFGFSXALGKTKTQSLVQNVAGRSFAVAD 437
QY 409 -----BEKTK---TIPAEVRKLTLEIPEM--DKOPDKQESLKITDIN 449
D 438 LAASLALFNKQAKIITATP-----VQLPNFGSGGSTPEQLPNPLNELIKIN 485

RESULT 4
AAT27522 PRELIMINARY; PRT; 491 AA.
AC AAT27522;
DT 24-MAY-2004 (TREMBlrel. 27, Created)
DT 24-MAY-2004 (TREMBlrel. 27, Last sequence update)
DE Unspecified sugar ABC transporter binding protein.
GN MNOB0360.
OS Mycoplasma mobile 163K.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma;
OX Mycoplasma mobile.
NCBI_TaxID=267748;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=163K;
RA Berg H.C.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=163K;
RA Birren B.W., Stange-Thomann N., Smith C., Decaprio D., Fisher S.,
Butler J., Calvo S., Ekin T., Fitzgerald M.G., Hafez N., Kodira C.D.,
Major J., Wang S., Wilkinson J., Nicol R., Nuebaum C.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=163K;
RA Jaffe J.D., Church G.M.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO17337; AAT27522.1; -
SQ SEQUENCE 491 AA; 52817 MW; 2D1C70768771B41 CRC64;
Query Match 20.3%; Score 467.5; DB 2; Length 491;
Best Local Similarity 29.7%; Pred. No. 1.5e-20;
Matches 141; Conservative 80; Mismatches 165; Indels 89; Gaps 18;
QY 43 NQKQITDVSIGISLVNERSEIIMAAKADKHGMLMAIVTAGGTNDNSFNQSSWEAI 102
DB 32 NTGLRVTDNOVFDLVAAREAFATQVANNSSLPNSKTLTLTAGGVNDLSFNQSSWEAI 91
QY 103 QQLGALTG--GEITSVDSSTA---ELEKYSILANTKNVWVLSGFQGDFTWKLPIE 157
DB 92 LEIGROTGRKGNFSFATTAAGTPOQROQDQAFPHHKFVVLTFQODGAFQWMLQGN 151
QY 158 NKQLEFKENIIIGIDMTDENVIPTGRYINLTYKTEBAGWLAGYANASFLAKKPSDPT 217
DB 152 NRAEFIRKQVIVADWTTLLEVPPOFISINRTQSSWIVGNAVAKFTSDHNNNN-- 209
QY 218 KBAIVVGGGISPAVNDPIAGYLAGIKAMLNKNS-----DKTKITT-DKIEI 264
DB 210 -RFTNFTGGGAFPEVTNFGNFGLOGILDFN--NSTFLEPGFTSTTDKKLSFTGDIINI 266
QY 265 NLGFDVQDTSTKERLEQIASKDKPST--LLAVAGPLTEIFSDLIANQND--RYLIGVDTQ 321
DB 267 NTGFAV----TPEAKAIQISVSGCTQVFPVAGSLTTLITVNSISQNSGQFVIGVDSQ 322
QY 322 SLVYTKTKNK-FTSILKNLGYSVFVLSLY-----TKKSN 358
DB 323 AKAFSPDLATLFFSSVEKNVAGTYAALASLYLGTGSTDPFNTGSSSRPIPTTEKNS 382
QY 359 RNLAGFFPKKSNVAVYIGIKDR-----FVADIATSLGNDKDLATEA-----ISEAKKEP 408
DB 383 SSLLP-----LANADITGVESTEPEVDFGFSKALGKTKTQSLVQANVGRSFEVAD 437
QY 409 -----BEKTK---TIPAEVKTULPIPEMP--DKQPDKQOESDKLTITDIN 449
DB 438 LAASLALFNKKNKIATIP-----VQIPNFSGGSTPQGINPLNELIKKIN 485

RESULT 6
Q9X775 PRELIMINARY; PRT; 465 AA.
AC Q9X775;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE P48 membrane lipoprotein precursor.
GN Name=P48;
OS Mycoplasma agalactiae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20002620; PubMed=10531294;
ID Q9E5B5 PRELIMINARY; PRT; 468 AA.
AC Q9E5B5;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE P48-like protein.
GN Name=P48;
OS Mycoplasma dovis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=28903;
RN [1]
RP SEQUENCE FROM N.A.
RA Rosati S., Alberti A., Robino P., Piteau M.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY57344; AAT64137.1; -
SQ SEQUENCE 468 AA; 51175 MW; AA0333F810BD2FEE CRC64;
Query Match 19.1%; Score 438.5; DB 2; Length 468;
Best Local Similarity 27.8%; Pred. No. 8.6e-19;
Matches 139; Conservative 93; Mismatches 179; Indels 89; Gaps 17;
QY 3 KKIKNKFTGLGLVFPISATATISAGCWDKETTKESADNQNQKQITDVSIGISLVNERK 62
DB 2 KKKKFFULGAA--PVLSPVLAASGCDYFKETEVDG-----VKTVTLSHIYSRK 52
QY 63 SEIMAAKADKHGMLMAIVTAGGTNDNSFNQSSWEAIQQLGALTGEITSV----- 116
DB 53 LKLRDGLTVDN---APRAAITDEGSVHDESFNQSWEAVHKISYELGLDKAQSNGKNL 109
QY 117 -----DSSTAELEKYSILANTKNVWVLSGFQGDFTWKLPIENKQLEFKENIIIG 171
DB 110 RNKRYEPRKGEKELASYSKNAIDSSFRYVILCGFTHKAL--YGLPEYIKIKIKNNIVFIT 167
QY 172 IDWTDTEENV-----IPTGRYINLTYKTEBAGWLAGYANASFLAKKPSDPTK 218
DB 168 VDDPIQDDASTGEPAAKAFYDKIGQGRLLVITPTKQAAVYAGBALDYSKIYKDNPEK 227
QY 219 RSAIVVGGGISPAVNDPIAGYLAGIKAMLNKNSDKTKITTDKIEINLGFVDQSTKER 278
DB 228 RTTGAFGIGIPWPSVDPIINFTPGIIDWNKEHEATKSLNNTIELTSP---TSGEPA 284
QY 279 LEQIASKDKPSTLLAVAGPLTEIFSDLIANQND--RYLIGVDTQSLVYTKKN-----K 331
DB 285 VAAINSVYKATASYPVAGSISDPAKEIKKLGKNNKFIIGVDADQ-----KVALKGRH 337
QY 332 FPTSILKNLGYSVFVLSLY-----TKKSNRN--AGEFGCKS-----ATVYIGIK 378
DB 338 IFTSVMKLIGQAVYNVLADLYSQGNSLSLOPGEIGKNGKGEAKVFGYGENESKIVGVA 397
QY 379 DRFVDIADTS--LEGNDKILATLAISEAKKEFEKTKTIPAEVKTUL--EIPK---MP 430
DB 398 -----TSGLDLSKNDLEIANKALBEATKYISK-----KALIKTLISQGLEAKKALG 444
QY 431 DKQPDKQOESLDKLTITDINK 450
DB 445 TKWPDQPADQFGKMINWLAK 464

RA Rosati S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M.,
RA Pletau M.;
RT "P48 major surface antigen of Mycoplasma agalactiae is homolog to a
RT major product of Mycoplasma fermentans and belongs to a selected family
RT of bacterial lipoproteins";
RL Infect. Immun. 67:6213-6216(1999).
DR EMBL; AJ132423; CAB43778.1;-.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
DR Lipoprotein; Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 465 AA; 51149 MW; 60AD5448CFE03C96 CRC64;

Query Match 18.1%; Score 417; DB 2; Length 465;
Best Local Similarity 28.4%; Pred. No. 1.8e-17;
Matches 141; Conservative 94; Mismatches 180; Indels 82; Gaps 22;

QY 3 KKIKMKFLGLVPLSAIATISAGCWDKETEKEKSNQNKQITDVSKISGL-VVER 61
DB 2 KKNKFYFLGNA---PVLSPVLVAASGDKYFKETEVDCVKTISTLAHITSRKGLREG 58
QY 62 KSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNQSSWEAIOQLGALTGEITVSSTA 121
DB 59 LTVENAPRA-----TFITDEGSVHDESFNQSWEAVHAKVSYELGLDKAQV-SGNK 107
QY 122 ELEBK-----YSSLANTNKVNVLSGFQHGDAFTRMKLIPEKNK-QLTEKNII 168
DB 108 NLRNKVEPEPKGQLLEAKNAIDSGFRYIVLCGFHQSL--VGLDENYIKIKIDNNII 164
QY 169 ILGIDWT---DTENV-----IPGRYINLYKTEBAGLAGYANASFLAKPEPDPDK 218
DB 165 FITVDPMFLTEPDDANVKTFIKKIGSHLVPIPTKQAAIYAGRALADYFQGVKXDEK 224
QY 219 RSAIVIGGGLSPAVTDFIAGLAGIKANLKNKSDKTKITTDKIBINLGPVQDSTKER 278
DB 225 RTIAGFGGIPMPAVSDFIAGTFQGIIDNNKEHPEAKTKSLNETIELNLF--TSGTPQA 281
QY 279 LEQIASKQRPSTLAVAGPL-TEIFSDI--IANQDRFLIGVDQSLVYTKRN----- 330
DB 282 TTAINSVVKAATASYPVAGSLSTIDRAKEIKLADK-DKTIIGVDADQ-----KNALKG 333
QY 331 KPFTSLKNLGYSVSVLSLYTKKNSRNL-AGFEFGKSAT-VYLGIDK----RFYDI 384
DB 334 RIFTSVMKLIGQAVYNILADLYSKGENQLDLPQFEIGKNGKTPVRYGDTEDKQYGV 393
QY 385 ADTS-LEGNDKKLATETAISEAKKEPEKTKTIPAEVAKTLEI-----DPMDDKQ 433
DB 394 ATSGILDDKNDKINAKLKDATAYVQK-KT-----EIQSKHDKDMETAKKALGAKFPBDP 448
QY 434 PDKQESLDKLTIDINK 450
DB 449 GGFQKQVMDWLASETRK 465

RESULT 7
052311 PRELIMINARY; PRT; 428 AA.
ID 052311;
AC 052311;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Membrane lipoprotein P48v.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG18;
RA Rawadi G., Dyer K., Dujancourt A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036106; AAB99740.1;-.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR InterPro; IPR008107; Mycoplasma_p48.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
DR Lipoprotein.
SQ SEQUENCE 428 AA; 47862 MW; D9506E817E330EAA CRC64;

Query Match 17.4%; Score 400.5; DB 2; Length 428;
Best Local Similarity 27.3%; Pred. No. 1.7e-16;
Matches 129; Conservative 95; Mismatches 175; Indels 73; Gaps 21;

QY 5 IKNNKFLGLVPLSAI-ATISAGCWDKETEKEKSNQNKQI--TDVSKISGLVNER 61
DB 1 MKSKIKALLGL-SPIALILPAVAISC-----KNNDSNISFKKEDISKYTTTANG 50
QY 62 KSEIMAAKADANKHFGMLMAIVTAGTVNDNSFNQSSWEAIOQLGALTGEITVSSTA 121
DB 51 KQVKNAL-----LKKPVLITDEGKIDKSFQSAFEALKAINKQITGEINSEPS-S 104
QY 122 ELEBGSLSANTNKVNVLSGFQHGDAFTRMKLIPEKNKQITDVSKISGLVNER 181
DB 105 NFEASINLSAGKIVLVNLFKQOSIKQY--IDAREELERQIKIGIDF-DITREY 161
QY 182 PTGRYINLYKTEBAGLAGYANASFLAKPEPDPTRSAIVIGGISPAVDFIAGYLA 241
DB 162 KW-FYSLQFNKISAFITGVAIASWLSQ---DESKRVVAFSGCAFPGVTTNKGFAK 216
QY 242 GIKAMNLKNSDKTKIT-TDKIELNLGFDVQDSTKERLQIAS-----KDKPST 290
DB 217 GILYNNQKH--KSYKFTHTSPVKLDSGF-----TAGKKNQTVINNVLSPPADVKYNPV 269
QY 291 ILAVAGPLTEIFSDIINQNDRLYIGVTDQSLVYTKKMFPSILKNLGYSVSVLS 350
DB 270 ILSVAGPAT--FEVYKLANQGQYIVIGVSDQGM--QDKRILTSVLKHKQAYETLLD 325
QY 351 L-----YTKKNSRNLAGEFGKSAATYVILGIDKRPVADIATSLGNDKKL-ATE 399
DB 326 LIEKEGVPYVVKDKADKKMSHFQTQ-----KEKIVGAENHPSNTERQAKINN 377
QY 400 AISEAKKEPEKTKTIPAEVAKTLEIPEMPD-KQPKQESLDKLTIDINK 450
DB 378 KIKKAIKPKF-----LPEDFVKYINSDKALDKGNKIDNVSERLAIIISAIK 425

RESULT 8
09RGX5 PRELIMINARY; PRT; 428 AA.
ID 09RGX5;
AC 09RGX5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Macrophage activating lipoprotein-404 precursor.
GN Name=malp;
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK5;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Different posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099211; AADI6395.1;-.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR InterPro; IPR008107; Mycoplasma_p48.

RP SEQUENCE FROM N.A.
RA Seva T., Matsumoto, M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; D64083; BAA23530.1; -
DR EMBL; AB026157; BAA77211.2; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR InterPro; IPR008107; Mycoplasma_D48.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LipROTEIN48.
SQ SEQUENCE 428 AA; 47862 MW; F43B078F21DAD6E CRC64;

Query March 17.2%; Score 394.5; DB 2; Length 428;
Best Local Similarity 26.8%; Pred. No. 3.9e-16;
Matches 126; Conservative 101; Mismatch 174; Indels 69; Gaps 20;

QY 5 IKNNKFLGLGVLPPLSAI-ATISAGCDKETTKEKSGADNQNKOITDVSKISGLWNERKS 63
Db 1 MKSKSKILLGL-SPIAIILPVAIVSCGNDES-----NISFEKDISKTYTTNANQKQ 52
QY 64 EIMAAKADANKHFGELNMAIYTAGTVNDNSFNQSGWEALIQGLALGTGETSVDSITAEI 123
Db 53 VVKNAEL-----LKLAKVLLITDEGKIDKSFNQSFAFALKAINKQTGIEINNVPS-SNF 106
QY 124 EGKSSSLANTNNKVVNLSGFQHGDAFIRWLKIPENKQLFTEKNIILIGIDPTENVIPT 183
Db 107 ESANISALSLAGHKIWLNVGFQHQSIKQY-IDAHREELERNQIKITIGIDF-DIETEKW 163
QY 184 GRVINTLVYKTEEAGWLGAVNANSEFLAKFPSPDPKRAIVIGGISPATVDFIAGYLAGI 243
Db 164 --FSLQFNKIKESAFITGVIAISLSQ---DESKRVVASGGGAPFGVTFNFGPAKI 218
QY 244 KAMNLKNSDKTKTI-TTDKIEINLGFVDVDTSTKERLEQIAS-----KKKPTLL 292
Db 219 LYYNQKH--KSSKIYHSTPVLDSCF-----TAGEKKNVIVNNVLSSTPADVKYNPVIL 271
QY 293 AVAGPLEIFESDILANDNDRYLVGDPDQSLVYKTKNKKPFSTLILKVLGVSFVSLADL 351
Db 272 SVAGPAT--PEYVALANKGGVIVGVDSDQGI--QDDRIITVSLKHKQAVVETLDDL 327
QY 352 -----YTKKSNSHNLGFEFGKGSATVYLGIDKRFVDIADTSLSEGNDKL-ATEAI 401
Db 328 LEKEGKRPVYVQKXKDKKMSHGTC-----KKKMGVAENHNSNTEQAKINKKI 379
QY 402 SEAKKEFEKTKTIPAEVRKTLIEIPMPD-KQPDQKQESLKLITDINK 450
Db 380 KEALIKMFE-----LPEDFVKYINDKALKKQCNKIDNVSERLEAIIAISAIK 425

RESULT 10
Q9RGX4 PRELIMINARY; PRT; 429 AA.
AC Q9RGX4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2004 (Tremblrel. 26, Last annotation update)
DE Macrophage activating lipoprotein-404 precursor.
GN Name=mlp;
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2115;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MT-2;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099212; AADI6396.1; -
DR GO; GO:0008289; F.lipid binding; IEA.
DR InterPro; IPR003760; Bmp.

DR InterPro; IPR008107; Mycoplasma_p48.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
KW Lipoprotein; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 429 macrophage activating lipoprotein-404.
SQ SEQUENCE 429 AA; 47933 MW; B6D08CF975AC3171 CRC64;
Query Match 17.1%; Score 394; DB 2; Length 429;
Best Local Similarity 27.0%; Pred. No. 4.2e-16;
Matches 127; Conservative 102; Mismatches 172; Indels 70; Gaps 21;
QY 5 IKNNKFLGLGVPLPSAI-ATISACQWDEKTEKESADNQNKOITDVSKISGLVNERKS 63
DB 1 MKSKSKILLGL-SPIAAILPAVAVSCGNNDSE-----NISFKKDIISKYTTTNNANQKQ 52
QY 64 EIMAKADANKHFGIMNAIYTAGTVNDNSFNQSWEAIOQLALTGGEITSVDSSTAE 123
DB 53 VVKNAEL-----LTKLPVLTIDEGKIDKSFNQSAFALAIKQTEIINNVEPS-SNF 106
QY 124 EGYTSSILANTKNVWVLSGFOHGAFTRWLKIPEKQOLFTEKNIIIGIDWTDTEENVIPT 183
DB 107 ESAVNSALSGHKIIVWNGFGHQSIKQY--IDAHRELEBNQIKIGIDF-DIETEKW 163
QY 184 GRYINLYTEEBAGWLAGYANASFLAKKFPSPDPTKRSATVIGGIGSPAVTDFIAGYLAGI 243
DB 164 --FYSLOFNIKESAFITGVALASWLSQ---DESKRVVAFGGAFPGVTFPNEGFAKGI 218
QY 244 KAMWLKNSDKTKI-TTDKIEINLGPVODSTYERLEQIAS-----KDKPSTLL 292
DB 219 LYVQKH--KSKLYHTSPVKLDSGF-----TAGKKNVTIINNVLSTPADVKXNPVIL 271
QY 293 AVAGPLLEIFSDI-IANQNDRYLIGVTDQSLVYTKTKNKFSTILKNLGYSVFSVSDL 351
DB 272 SVAGPAT--PEYVRLAANKQGYIVGVDSDQGM--QDKDRLISVLGHITKQAVYETLLDL 327
QY 352 -----YTKSNSRNLAGFEFGKSAIVYLGIKORFVDIADTSLSEGNKKL-ATEA 400
DB 328 ILKEEGYKPYVVDKADKWSHFQ-----KEKWIQVAFNHSNTEBQAKINNK 379
QY 401 ISEAKKEFEEKTKTIPAEVRYKTLIEPMPD-KOPDQOESLDLITDINK 450
DB 380 IKEAIKMFKE---LPEDFVKYINSKALKDGNKIDNVSEKLEAIIISAINK 426
RESULT 11
ID Q9RGX6 PRELIMINARY; PRT; 429 AA.
AC Q9RGX6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Macrophage activating lipoprotein-404 precursor.
GN Name=malp;
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCB1_TaxID=2115;
RN NC1_TaxID=2115;
RP SEQUENCE FROM N.A.
RC STRAIN=Incognitus;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099210; AAD16394.1;-
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
KW Lipoprotein; Signal.

FT SIGNAL 1 24 Potential.
FT CHAIN 25 429 macrophage activating lipoprotein-404.
SQ SEQUENCE 429 AA; 47961 MW; B6D09A8812AC3171 CRC64;
Query Match 17.1%; Score 394; DB 2; Length 429;
Best Local Similarity 27.0%; Pred. No. 4.2e-16;
Matches 128; Conservative 101; Mismatches 169; Indels 76; Gaps 21;
QY 5 IKNNKFLGLGVPLPSAI-ATISACQWDEKTEKESADNQNKOITDVSKISGLVNERKS 63
DB 1 MKSKSKILLGL-SPIAAILPAVAVSCGNNDSE-----NISFKKDIISKYTTTNNANQKQ 52
QY 64 EIMAKADANKHFGIMNAIYTAGTVNDNSFNQSWEAIOQLALTGGEITSVDSSTAE 123
DB 53 VVKNAEL-----LTKLPVLTIDEGKIDKSFNQSAFALAIKQTEIINNVEPS-SNF 106
QY 124 EGYTSSILANTKNVWVLSGFOHGAFTRWLKIPEKQOLFTEKNIIIGIDWTDTEENVIPT 183
DB 107 ESAVNSALSGHKIIVWNGFGHQSIKQY--IDAHRELEBNQIKIGIDF-DIETEKW 163
QY 184 GRYINLYTEEBAGWLAGYANASFLAKKFPSPDPTKRSATVIGGIGSPAVTDFIAGYLAGI 243
DB 164 --FYSLOFNIKESAFITGVALASWLSQ---DESKRVVAFGGAFPGVTFPNEGFAKGI 218
QY 244 KAMWLKNSDKTKI-TTDKIEINLGPVODSTYERLEQIAS-----KDKPSTLL 292
DB 219 LYVQKH--KSKLYHTSPVKLDSGF-----TAGKKNVTIINNVLSTPADVKXNPVIL 271
QY 293 AVAGPLLEIFSDI-IANQNDRYLIGVTDQSLVYTKTKNKFSTILKNLGYSVFSVSDL 351
DB 272 SVAGPAT--PEYVRLAANKQGYIVGVDSDQGM--QDKDRLISVLGHITKQAVYETLLDL 327
QY 352 -----YTKSNSRNLAGFEFGKSAIVYLGIKORFVDIADTSLSEGNKKL-ATEA 400
DB 328 ILKEEGYKPYVVDKADKWSHFQ-----KEKWIQVAFNHSNTEBQAKINNK 379
QY 401 ISEAKKEF---EKTCTIPAEVRYKTLIEPMPD-KOPDQOESLDLITDINK 450
DB 380 IKEAIKMFKELPEDFVKYINSKALKDGNKIDNVSEKLEAIIISAINK 426
RESULT 12
ID Q9RGX7 PRELIMINARY; PRT; 428 AA.
AC Q9RGX7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Macrophage activating lipoprotein-404 precursor.
GN Name=malp;
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCB1_TaxID=2115;
RN NC1_TaxID=2115;
RP SEQUENCE FROM N.A.
RC STRAIN=II-29/1;
RX MEDLINE=99115554; PubMed=9916088;
RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans.";
RL Infect. Immun. 67:760-771(1999).
DR EMBL; AF099209; AAD16393.1;-
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PR01733; LIPOPROTEIN48.
KW Lipoprotein; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 428 macrophage activating lipoprotein-404.
SQ SEQUENCE 428 AA; 47960 MW; D6B5C0BA965CFB1 CRC64;

Query Match 17.1%; Score 393.5; DB 2; Length 428;
 Best Local Similarity 26.8%; Pred. No. 4.5e-16;
 Matches 127; Conservative 97; Mismatches 178; Indels 71; Gaps 21;

QY 1 MKKKIKNNKFLGLVPLSAIATIISAGCWDKETTKEKSDADNQNKOITDVSKISGLVNERK 60
 DB 1 MKKSKK--ILLGLSPIAAILPAVAVSCG-----NNDESINISFKKDIKITTNN--ANE 50

QY 61 RKSEIMAAKADANKHFLGNMAIYTAGTVNDSPNOSWEAIQOLGALTGEITSDVSST 120
 DB 51 KQ--VVKNAELK--LKPVLITDEKIDKSPNOSAFELKAIKQGTGIEINVSPE- 103

QY 121 ALEEGKSSILANTKNVWVLSGFOHDAFTRWLKIPEKOLFTEKNIILIGIMTDEENV 180
 DB 104 SNRESAYNSALSGHKIWLNGFRHQOSIKQY--IDARELERNOIKITIGID-DIETE 160

QY 181 IPTGRYINLTYTEEAGMLAGVANASFLAKKPSDPTKSAIYIGGISPAVTFIAGYL 240
 DB 161 YKW--FYSLOFNIKESAFITGVAIASWLSQ---DESKRVVASFGGAFPGVTFNNGFA 215

QY 241 AGIKAMLNKSDKTKI--TTDKIEINLGFVDVDTSTERLEQIAS-----KDKPS 289
 DB 216 KGLIYVQKH--KSSKIYHTSPVKLDSGF-----TAGKKNATVINNVLSSTPADVKYNPH 268

QY 290 TLAVAGPLTEIFSDIANGNDRYLIGVDTPDOSLVYTKTKNKFSTILKNLGYSVFVLS 349
 DB 269 VILSVAGPAT--PEYVLANKGVYIGVSDQGM--QDKDRILTSVLKIKQAVETL 324

QY 350 DL-----YTKSNSHNLGAFEGKKSATVYIGIKDFVDIADTSLGNDKUL-AT 398
 DB 325 DLLEKEGYKPYVVDKADKKMSHFGTQ-----KEKMGVAVENHFSNTEQAKIN 376

QY 399 EATSEAKKEFEKTKITPAEEVKTLEIPEMPD-KQPDQOESLDKLITDINK 450
 DB 377 NKIKKAIKMKKE---LPEDFVKYINSDKALKDNKIDNVSERLEAIISAINK 425

RESULT 13
 QYR3X3 PRELIMINARY; PRT; 428 AA.

AC Q9R3X3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Macrophage activating lipoprotein-404 precursor.
 GN Name=malp;
 OS Mycoplasma fermentans.
 OC Bacteria; Firmicutes; Molluscites; Mycoplasmataceae; Mycoplasma.
 NCBI_Taxid=2115;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K7;
 RX MEDLINE=99115554; PubMed=9916088;
 RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
 RT "Differential posttranslational processing confers intraspecies
 variation of a major surface lipoprotein and a macrophage-activating
 lipopeptide of Mycoplasma fermentans.";
 RL Infect. Immun. 67:760-771(1999).
 DR EMBL; AF099213; AAD16397.1; -;
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR InterPro; IPR003760; Bmp.
 DR InterPro; IPR008107; Mycoplasma_p48.
 DR Pfam; PF02608; Bmp; 1.
 DR PRINTS; PR01733; LIPOPROTEIN48.
 KM Lipoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN 25 24 Potential.
 SO SEQUENCE 428 AA; 47875 MW; C7A536B409A60132 CRC64;

Query Match 16.9%; Score 389.5; DB 2; Length 428;
 Best Local Similarity 26.8%; Pred. No. 8e-16;
 Matches 126; Conservative 99; Mismatches 176; Indels 69; Gaps 20;

QY 5 IKNNKFLGLVPLSAI-ATISAGCWDKETTKEKSDADNQNKOITDVSKISGLVNERK 63
 DB 1 MKSKKILDL-SPIAAILPAVAVSCGNDES-----NISFKKDIKITTNNANQKQ 52

QY 64 EIMAAKADANKHFLGNMAIYTAGTVNDSPNOSWEAIQOLGALTGEITSDVSSTAE 123
 DB 53 VVKNAEL-----LKLKVLITDESKIDKSPNOSAFELKAIKQGTGIEINVSPE-SNF 106

QY 124 EGYKSSILANTKNVWVLSGFOHDAFTRWLKIPEKOLFTEKNIILIGIMTDEENV 183
 DB 107 ESAYNSALSGHKIWLNGFRHQOSIKQY--IDARELERNOIKITIGID-DIETRYK 163

QY 184 GRYINLTYTEEAGMLAGVANASFLAKKPSDPTKSAIYIGGISPAVTFIAGYLAGI 243
 DB 164 --FYSLOFNIKESAFITGVAIASWLSQ---DESKRVVASFGGAFPGVTFNNGFAKGI 218

QY 244 KAMNLKNSDKTKI--TTDKIEINLGFVDVDTSTERLEQIAS-----KDKPSTLL 292
 DB 219 LYYVQKH--KSSKIYHTSPVKLDSGF-----TAGKKNATVINNVLSSTPADVKYNPHVIL 271

QY 293 AVAGPLTEIFSDIANGNDRYLIGVDTPDOSLVYTKTKNKFSTILKNLGYSVFVLS 351
 DB 272 SVAGPAT--PEYVLANKGVYIGVSDQGM--QDKDRILTSVLKIKQAVETL 327

QY 352 -----YTKSNSHNLGAFEGKKSATVYIGIKDFVDIADTSLGNDKUL-ATEAI 401
 DB 328 LEKEGYKPYVVDKADKKMSHFGTQ-----KEKMGVAVENHFSNTEQAKIN 379

QY 402 SEAKKEFEKTKITPAEEVKTLEIPEMPD-KQPDQOESLDKLITDINK 450
 DB 380 KEAIKMKKE---LPEDFVKYINSDKALKDNKIDNVSERLEAIISAINK 425

RESULT 14
 QYR3N6 PRELIMINARY; PRT; 428 AA.

AC Q9R3N6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 03-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Lipoprotein MALP-404 precursor (Macrophage activating lipoprotein-404 precursor).
 GN Name=malp; Synonym=malp;
 OS Mycoplasma fermentans.
 OC Bacteria; Firmicutes; Molluscites; Mycoplasmataceae; Mycoplasma.
 NCBI_Taxid=2115;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PG18, M39A, and M70B;
 RX MEDLINE=99115554; PubMed=9916088;
 RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
 RT "Differential posttranslational processing confers intraspecies
 variation of a major surface lipoprotein and a macrophage-activating
 lipopeptide of Mycoplasma fermentans.";
 RL Infect. Immun. 67:760-771(1999).
 DR EMBL; AF099214; AAD16398.1; -;
 DR EMBL; AF099215; AAD16399.1; -;
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR InterPro; IPR003760; Bmp.
 DR InterPro; IPR008107; Mycoplasma_p48.
 DR Pfam; PF02608; Bmp; 1.
 DR PRINTS; PR01733; LIPOPROTEIN48.
 KM Lipoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN 25 38 Potential.
 SO SEQUENCE 428 AA; 47863 MW; 59808324E218F03C CRC64;

Query Match 16.9%; Score 387.5; DB 2; Length 428;
 Best Local Similarity 26.6%; Pred. No. 1.1e-15;
 Matches 125; Conservative 101; Mismatches 175; Indels 69; Gaps 20;

QY 5 IKNNKFLGLVPLSAI-ATISAGCWDKETTKEKSDADNQNKOITDVSKISGLVNERK 63

```

Db      1 MKSKSKILLGL-SP1AALPAVAVSCGNDES-----NISFKEDSKYTTTNNANGKQ 52
Qy      64 EIMAKADANGHFGIINMAITVAGTVNDNSFNQSSWEAIIQOLGALTGEIITSVDSFAEL 123
Db      53 VKRAHEL-----LKLKPVLTIDEGKIDDKSPNOSAFELKAIKNOTGIEINSVERP-SNF 106
Qy      124 EGKYSILANTNKVNWVLSGFQHGDAFTRWLKI PENKQLFTEKNIIIGIDMTDTENVIPT 183
Db      107 ESAVNSALISAGHKIIVLNGFGHQSIKQY--IDAHELEBERNQIKIIGIDE-DIETERYKW 163
Qy      184 GRVYNLYTTEAGMLAGYANASFLAKKFPSPDPTKRSALIVGGISPAVTFPIAGYLAGI 243
Db      164 --FYSLQFNIKESAFTTGIALASWLSQ--DESKRVASFGVAFPGVTFNFGFAGKI 218
Qy      244 KAWNLKNSDKKTKI-TTDKIEINLGFDPVDTSTKERLEQIAS-----KDKPSTLL 292
Db      219 LYYNQKH--KSKSYIHSFVLDLSGF-----TAGEKMTVINNVLSSTPAAVTKTNPHVYL 271
Qy      293 AVAGPLTEIFSDIIANQNDRYLIGVDTQSLVYTKKXKFFTSILKNLGYSVFVSLSDL- 351
Db      272 SVAGPAT--FETVLANKGQYVIGVSDQGMII--QDKORILTSVLKHIKQAVYETLIDLII 327
Qy      352 -----YTKKNSRMLAGFEFGKKSATVYLGIKDRFVIDADTSLEGNDKUL-ATBAI 401
Db      328 LEKEEGKYPYVVKADKADKMSHFQO-----KEKMGVAENHPSNTEQAKINNKI 379
Qy      402 SEAKKEFEKTKTTPAEERKTLIEPEMD-KOPDKQOESLDKLTITDINK 450
Db      380 KEALIMFRE---LPEDFVKIINSDKALDKGNKIDNVSERLEAIISAINK 425

```

RESULT 15

```

ID      08RCH2      PRELIMINARY;      PRT;      352 AA.
AC      08RCH2;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT      01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE      Surface lipoprotein.
GN      Name=Med; OrderedLocusNames=TTE0457;
OS      Thermoaerobacter tengcongensis.
OC      Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;
OC      Thermoaerobacteriaceae; Thermoaerobacter.
OX      NCBI_TaxID=119072;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MB4;
RX      MEDLINE=21992816; PubMed=11997336;
RA      Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA      Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA      Tan H., Chen R., Wang J., Yu J., Yang H.;
RT      "A complete sequence of the T. tengcongensis genome.";
RL      Genome Res. 12:689-700(2002).
DR      EMBL; AE013017; FANM23740.1;
DR      GO; GO:0008289; F:lipid binding; IEA.
DR      InterPro; IPR003760; Bmp.
DR      Pfam; PF02608; Bmp; 1.
KW      Complete proteome; Lipoprotein.
SQ      SEQUENCE 352 AA; 38052 MW; DCA7990EA271A23B CRC64;

```

Query Match 10.7%; Score 247; DB 2; Length 352;
 Best Local Similarity 25.7%; Pred. No. 3.7e-07;
 Matches 105; Conservative 62; Mismatches 161; Indels 80; Gaps 17;

```

Qy      6 KMNKFLGLGVFPISATATISAGCMDEKTTKEKSADNQNKOITDVSKISLVNKRSEI 65
Db      3 KMRVVLASLLTLIV-ALSVLISGSSK--TKQEST--QPOQTTEANK-----44
Qy      66 MAATADANKHFGIINMAITVAGTVNDNSFNQSSWEAIIQOLGALTGEIITSVDS-STAELE 124
Db      45 -----NKNF--KVGIVTVDVGINDRSFNQMAIEGLQRAKEIGVTVNVVQSOMPTDV 95

```

```

Qy      125 GKYSILANTNKVNWVLSGFQHGDAFTRWLKI PENKQLFTEKNIIIGIDMTDTENVIPTG 184
Db      96 PNLTFNFAQCGYDLVIVSGFMMHDA-----TEEVSQKFPYTKFLLIDSEITDRPNV----- 145
Qy      185 RYINLYTTEAGMLAGYANASFLAKKFPSPDPTKRSALIV--IGGISPAVTFPIAGYLAGI 242
Db      146 --ASAMFEQEVGYLAG--ALAGLVEKEKYGKYGKGINIIGAVGGMQIPPDRTFAGYQGG 201
Qy      243 IKAWNLKNSDKKTKITTDKIEINLGFDPVDTSTKERLEQIASDKDPSTLLAVAGPLTEIF 302
Db      202 AKAVN-----PDIKILINVTNNFNPDPAGKQWALNQISQGAELIFQVAGGTGB-- 249
Qy      303 SDIT--ANQNDRYLIGVDTQSLVYTKKXKFFTSILKNLGYSVFVSLSD-----350
Db      250 -GVYKAAQERKNLYAIGVDADQSYL--APDNVLTSAVKRYDVAAYVDVYKALANGNFKSGI 305
Qy      351 LYTKKNSRMLAGFEFGKKSATVYLGIKDRP---VDIADTSLEGNDK 394
Db      306 MYFDLKNN---GVGLKINKVDVQSIIDQVNOALAKDIIIDGKIQVSDK 349

```

Search completed: December 18, 2004, 01:26:31
 Job time : 131.457 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:06:47 ; Search time 106.476 Seconds
(without alignment)
1425.134 Million cell updates/sec

Title: US-09-676-249D-4

Perfect score: 2155
Sequence: 1 MMDEKTTKEKSAADNQNQI.....KQPKQOESIDKLTIDINL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	423	5	AAO15862 Mutant My
2	2134	99.0	451	4	AAU01859 Mycoplasma
3	2134	99.0	451	5	AAO15861 Mycoplasma
4	2120	98.4	423	4	AAU01860 Mycoplasma
5	702	32.6	453	4	AAU01860 Mycoplasma
6	693.5	32.2	457	5	AAO15867 Mycoplasma
7	692.5	32.1	432	2	AAU01867 Mycoplasma
8	642.5	29.8	448	1	AAU01867 Mycoplasma
9	379.5	17.6	428	6	AAU01867 Mycoplasma
10	379.5	17.6	428	6	AAU01867 Mycoplasma
11	378.5	17.6	404	7	AAU01867 Mycoplasma
12	378.5	17.6	428	7	AAU01867 Mycoplasma
13	375.5	17.4	404	7	AAU01867 Mycoplasma
14	375.5	17.4	428	7	AAU01867 Mycoplasma
15	372	17.3	429	2	AAU01867 Mycoplasma
16	220.5	10.2	350	5	ABP25843 Streptococ
17	220.5	10.2	350	5	ABP25843 Streptococ
18	217	10.1	349	5	ABP27752 Streptococ
19	216.5	10.0	330	5	ABG66916 Streptococ
20	216.5	10.0	330	5	ABG66915 Streptococ
21	213.5	9.9	330	5	ABG66917 Streptococ
22	206	9.6	351	3	AAU01863 Streptococ
23	204	9.5	363	8	AAU01863 Streptococ
24	203.5	9.4	350	5	AAU01863 Streptococ
25	202	9.4	328	2	AAU01863 Streptococ

26	202	9.4	328	5	ABP54560 S. pneumo
27	202	9.4	328	7	ADCA5089 S. pneumo
28	202	9.4	350	6	ABU01234 S. pneumo
29	202	9.4	374	7	ADCA5457 E. faeciu
30	195.5	9.1	347	5	ABP10886 Streptococ
31	195.5	9.1	347	5	ABP28678 Streptococ
32	195.5	9.1	370	5	ABP25842 Streptococ
33	179.5	8.3	366	6	ADBO9398 Streptococ
34	178	8.3	339	2	AAU00049 Enterococ
35	178	8.3	339	5	ABP43268 E. faecali
36	178	8.3	339	6	ABU88296 E. faecal
37	178	8.3	339	6	ABU13547 Enterococ
38	178	8.3	341	2	AAU61758 B. burgdo
39	178	8.3	361	2	AAU00048 Enterococ
40	178	8.3	361	5	ABP43267 E. faecali
41	178	8.3	361	6	ABU88295 E. faecal
42	178	8.3	361	6	ABU13546 Enterococ
43	178	8.3	375	7	ADH88120 Enterococ
44	174	8.1	347	5	ABP29943 Streptococ
45	174	8.1	347	5	ABG66913 Streptococ

ALIGNMENTS

RESULT 1	AAO15862	standard; protein; 423 AA.
ID	AAO15862	
XX	AAO15862;	
AC	AAO15862;	
DT	16-JAN-2003	(first entry)
XX		
DE	Mutant Mycoplasma hyopneumoniae mhp3 protein.	
XX		
KW	Apoptein antigen; enzootic mycoplasma pneumoniae; mhp3; mutant; mhp3; vaccine; Mycoplasma hyopneumoniae infection.	
XX		
OS	Mycoplasma hyopneumoniae.	
XX	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FT	Misc-difference 70	/note= "Encoded by RGT"
FT	Misc-difference 422	
FT	Misc-difference 423	/note= "Wild type Lys replaced by Asn"
FT	Misc-difference 423	/note= "Wild type Asn replaced by Leu"
XX		
PN	EP1245677-A1.	
XX		
PD	02-OCT-2002.	
XX		
XX	30-MAR-2001; 2001EP-00303030.	
XX		
XX	30-MAR-2001; 2001EP-00303030.	
XX		
PA	(PRIZ) PRIZER PROD INC.	
XX		
PI	King KW, Madura RA, Rosey EL;	
XX		
DR	WPI; 2002-742716/81.	
XX		
XX	N-PSDB; AAL50633.	
XX		
XX	Novel apoprotein antigens encoded by Mycoplasma hyopneumoniae for use in	
XX	vaccines to prevent and treat diseases caused by infection with	
XX	Mycoplasma hyopneumoniae in animals, especially pigs.	
XX	Claim 4; Page 20-21, 38pp; English.	
XX		
CC	The invention comprises the amino acid and coding sequences of Mycoplasma	
CC	hyopneumoniae mhp3 proteins, the invention also comprises novel	
CC	apoprotein antigens encoded by the M. hyopneumoniae mhp3 gene. M.	


```

RESULT 3
AA015861
ID AA015861 standard; protein; 451 AA.
XX
AC AA015861;
XX
DT 16-JAN-2003 (first entry)
XX
DE Mycoplasma hyopneumoniae mhp3 protein.
XX
KM Apoptotectin antigen; enzootic mycoplasma pneumoniae; mhp3; vaccine;
KW Mycoplasma hyopneumoniae infection.
XX
OS Mycoplasma hyopneumoniae.
XX
FH Key Location/Qualifiers
FT Misc-difference 7 /note= "Encoded by TGA"
FT Misc-difference 99 /note= "Encoded by TGA"
FT Misc-difference 138 /note= "Encoded by TGA"
FT Misc-difference 152 /note= "Encoded by TGA"
FT Misc-difference 174 /note= "Encoded by TGA"
FT Misc-difference 198 /note= "Encoded by TGA"
FT Misc-difference 246 /note= "Encoded by TGA"
FT Misc-difference 246 /note= "Encoded by TGA"
XX
PN EP1245677-A1.
XX
PD 02-OCT-2002.
XX
PF 30-MAR-2001; 2001EP-00303030.
XX
PR 30-MAR-2001; 2001EP-00303030.
XX
PA (PF1Z ) PFIZER PROD INC.
XX
PI King KM, Madura RA, Rosey EL;
XX
DR WPI; 2002-742716/81.
XX
N-PSDB; AAL5632.
XX
PT Novel apoprotein antigens encoded by Mycoplasma hyopneumoniae for use in
PT vaccines to prevent and treat diseases caused by infection with
PT Mycoplasma hyopneumoniae in animals, especially pigs.
XX
PS Claim 11; Fig 1; 38pp; English.
XX
XX The invention comprises the amino acid and coding sequences of Mycoplasma
CC hyopneumoniae mhp3 proteins, the invention also comprises novel
CC apoprotein antigens encoded by the M. hyopneumoniae mhp3 gene. M.
CC hyopneumoniae is a bacterial pathogen that causes enzootic mycoplasma
CC pneumonia in pigs. The mhp3 genes, proteins and apoprotein antigens of
CC the invention are useful in the manufacture of a vaccine for treating/
CC preventing a disease or disorder caused by infection with M.
CC hyopneumoniae in an animal, especially a pig. The present amino acid
CC sequence represents a Mycoplasma hyopneumoniae mhp3 protein of the
CC invention
XX
SQ Sequence 451 AA;
Query Match 99.0%; Score 2134; DB 5; Length 451;
Best Local Similarity 99.8%; Pred. No. 2,2e-162;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 WDKEETKEKSDNONKOITDVSKISGLVNERKSEIMAAKADANKHFGLNNAVITAGGTG 61
DB 30 WDKEETKEKSDNONKOITDVSKISGLVNERKSEIMAAKADANKHFGLNNAVITAGGTG 89

```

```

OY 62 NDNSFNQSGWEALIOQLGALTGGEITTSVDSSSTAIEGKYSSLANTKNVWVLSCFQHGDAF 121
DB 90 NDNSFNQSGWEALIOQLGALTGGEITTSVDSSSTAIEGKYSSLANTKNVWVLSCFQHGDAF 149
OY 122 TRWLKIPENKQLFTEKNIIILGIDWTDTEENVIPGRYINLTYYKTEAGWLAGYANASFLA 181
DB 150 TRWLKIPENKQLFTEKNIIILGIDWTDTEENVIPGRYINLTYYKTEAGWLAGYANASFLA 209
OY 182 KKPSDPTKRSALVIGGSIAPAVTDFIAGYLAGIKANLNKSDKKTITTDKIEINIGFD 241
DB 210 KKPSDPTKRSALVIGGSIAPAVTDFIAGYLAGIKANLNKSDKKTITTDKIEINIGFD 269
OY 242 VQDSTKERLEQJLAKSKPSTLLAVAGPLTEIFSDIIANONDRYLIGVTDQSLVYTKTK 301
DB 270 VQDSTKERLEQJLAKSKPSTLLAVAGPLTEIFSDIIANONDRYLIGVTDQSLVYTKTK 329
OY 302 NKFEFTSLIKNLGVSVFSLDLYTKKSNRNLAGFEFGKKSATVYLGIKORFVDIADTSL 361
DB 330 NKFEFTSLIKNLGVSVFSLDLYTKKSNRNLAGFEFGKKSATVYLGIKORFVDIADTSL 389
OY 362 EGNDKRLATEAISAKKEFEKTKTTPAEVRKTLTEPMPDQPKDQGSIDLKLTIDIN 421
DB 390 EGNDKRLATEAISAKKEFEKTKTTPAEVRKTLTEPMPDQPKDQGSIDLKLTIDIN 449

RESULT 4
AA01860
ID AA01860 standard; protein; 423 AA.
XX
AC AA01860;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycoplasma hyopneumoniae recombinant MHP3 antigen.
XX
KM MHP3; antigen; vaccine; enzootic mycoplasma pneumoniae; mutant; antibody;
KW immunosassay; immunotherapy; anti-idiotypic antibody; mutein.
XX
OS Mycoplasma hyopneumoniae.
XX
FH Key Location/Qualifiers
FT Misc-difference 70 /note= "Encoded by RGA"
FT Misc-difference 417.422 /note= "Encoded by AATTAACGATAT"
XX
PN EP1090995-A2.
XX
PD 11-APR-2001.
XX
PF 26-SEP-2000; 2000EP-00308421.
XX
PR 29-SEP-1999; 99US-0156602P.
XX
PA (PF1Z ) PFIZER PROD INC.
XX
PI King KM, Madura RA, Rosey EL;
XX
DR WPI; 2001-309781/33.
XX
N-PSDB; AAS03286.
XX
PT New apoprotein antigens encoded by mhp3 gene from Mycoplasma
PT hyopneumoniae useful as a vaccine for treating or preventing diseases
PT caused by Mycoplasma hyopneumoniae.
XX
PS Claim 4; Page 19-21; 38pp; English.
XX
XX The sequence represents Mycoplasma hyopneumoniae a recombinant MHP3
CC antigen lacking the first 28 amino acids (the putative signal sequence).
CC MHP3 antigen and its fragments are useful in manufacturing a vaccine for
CC treating or preventing a disease or disorder in an animal, especially
CC pig, caused by M. hyopneumoniae infection e.g. enzootic mycoplasma

```

CC pneumonia. The mbp3-encoded proteins may be used as immunogens to
CC generate antibodies which immunospecifically bind such an immunogen. The
CC antibodies generated against the antigen are useful in diagnostic
CC immunoassays, passive immunotherapy and generation of anti-idiotypic
CC antibodies. Mb3 proteins may also be used in immunoassays, e.g. to
CC detect or measure in a biological sample from a vaccinated or potentially
CC infected test animal the presence of antibodies to the antigen, and thus
CC to monitor the immune response and/or to diagnose infection of the animal

XX
XX Sequence 423 AA;

Query Match 98.4%; Score 2120; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 2,7e-161; Indels 0; Gaps 0;
Matches 416; Conservative 0; Mismatches 0;

QY 1 MMDKETTKEKSAQNOKQITDVSKISGLVNERKSEIIMAAKADANKHGLNMAIVTAGCT 60
DB 1 MMDKETTKEKSAQNOKQITDVSKISGLVNERKSEIIMAAKADANKHGLNMAIVTAGCT 60
QY 61 VNDNSFNQSGWEAIIQQLGALTGTGETTSDVSSSTAELGKYSLSANTNNKVVWLSGFGHDA 120
DB 61 VNDNSFNQSGWEAIIQQLGALTGTGETTSDVSSSTAELGKYSLSANTNNKVVWLSGFGHDA 120
QY 121 FTRWLKIPENKQLFTEKRIIILGIDMTENVIPGRYINLTYTEEAGWLAGYANASFL 180
DB 121 FTRWLKIPENKQLFTEKRIIILGIDMTENVIPGRYINLTYTEEAGWLAGYANASFL 180
QY 181 AKKPPSPDKSAIVIGGSIPTVDFIAGYLAGIKANLKNSDKTKITTDKIEINLGF 240
DB 181 AKKPPSPDKSAIVIGGSIPTVDFIAGYLAGIKANLKNSDKTKITTDKIEINLGF 240
QY 241 DVQDTSTYERLEQIASKDKPSTLLAVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKT 300
DB 241 DVQDTSTYERLEQIASKDKPSTLLAVAGPLTEIFSDIIANONDRYLIGVDTQSLVYTKT 300
QY 301 KNKFTSTILKNLGYSVFSVLSDLTYKKSNSRLAGFERGKSATVYIGIKRPFDIADTS 360
DB 301 KNKFTSTILKNLGYSVFSVLSDLTYKKSNSRLAGFERGKSATVYIGIKRPFDIADTS 360
QY 361 LEGNDKILATEAISEAKKEFEKTKTIPAEVRKTLIEIPMPDKQPDQKQESLDEL 416
DB 361 LEGNDKILATEAISEAKKEFEKTKTIPAEVRKTLIEIPMPDKQPDQKQESLDEL 416

RESULT 5
AAR40856
ID AAR40856 standard; protein, 453 AA.
XX
AC AAR40856;
XX
DT 25-MAR-2003 (revised)
DT 07-MAR-1994 (first entry)
XX
DE 43kd regression associated antigen.
XX
KM Regression associated antigen; tumour; immunotherapy;
XX anti-idiotypic antibodies; antibodies; tumour regression.
XX
OS Mycoplasma hyorhinis.
XX
FH Key Location/Qualifiers
FT Misc-difference 80
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT hyorhinis."
FT Misc-difference 124
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT hyorhinis."
FT Misc-difference 165
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT hyorhinis."
FT Misc-difference 344
FT /note= "Tryptophan encoded by TGA, normal in Mycoplasma
FT hyorhinis."

XX
PN US5242823-A.
XX
PD 07-SEP-1993.
XX
PP 02-OCT-1992; 92US-00956546.
XX
XX 07-MAR-1986; 86US-00837494.
PR 16-SEP-1987; 87US-00097910.
PR 11-DEC-1987; 87US-00131815.
PR 04-JAN-1988; 88US-00138923.
PR 16-MAR-1990; 90US-00474730.
XX
XX (ITGE-) INT GENETIC ENG INC.
XX
XX Fareed GC, Sen A, Ghosh-Dastidar P, Jar-How L;
XX WPI: 1993-295229/37.
XX DR N-PSDB; AA047816.
XX
XX DNA encoding a regression-associated antigen from M. hyorhinis - is used
XX to obtain prods. for diagnosis, localisation and therapy of tumours.
XX
XX Disclosure; Fig 3; 40pp; English.
XX
XX Regression associated antigens (RAA's) are identified in material from
XX neoplastic cells by their immunological reactivity with regression
XX associated antibodies from the serum of patients diagnosed as undergoing
XX regression of a tumour. RAA's can be used for tumour immunotherapy and
XX for producing and purifying antibodies which can be used for tumour
XX diagnosis, localisation and therapy. The antibodies can also be used for
XX the production of anti-idiotypic antibodies which can also be used in
XX immunotherapy. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 453 AA;

Query Match 32.6%; Score 702; DB 2; Length 453;
Best Local Similarity 38.8%; Pred. No. 1,7e-47; Indels 40; Gaps 14;
Matches 170; Conservative 82; Mismatches 146;

QY 5 ETTKEESADNOKQITDVSKISGLVNERKSEIIMAAKADANKHGLNMAIVTAGTVNDN 64
DB 28 ETTKEESKIT-----RIID-----NSFVDRQAIEKA-----KNDFRTVLLTAGTVQDK 73
QY 65 SFNQGWEAI---QQLGALTG-GETTSVDSSTAELGKYSLSANTNNKVVWLSGFGHGD 119
DB 74 SFNQSWEAVLEHNDQIEKTTNLDVRSQETNNQSELIGKYNFLNGKNVWILTGFOGQ 133
QY 120 AFTRWLKIPEH-----KQLTEKRIIILGIDW---TTEENVIPGRYINLTYTEEAGWL 171
DB 134 EFPKFLKQDTSNGKRYSDLAEEKVITVAVDWLSKEDKDLIKAGHFIISLYKTEAGFI 193
QY 172 AGYANASFLAKKPPSPDKSAIVIGGSIPTVDFIAGYLAGIKANLKNSDKTKITTT 231
DB 194 AGYASSKFLAYKFPNDAKKTIAPFGGSHGAGVDFIAGYLAGIAKANNNDPTAKVITSD 253
QY 232 DKIEINLGFVQDTSTYERLEQIASKDKPSTLLAVAGPLTEIFSDII--ANONDRYLIGV 289
DB 254 NNINIDTGF-ISNDKTAFTINGIVNKS--SLVPEVGSLSNVDAIKKSKDKTKYDILGV 310
QY 290 DTDQSLVYTKTKKKEFFSILKNLGYSVFSVLSDLTYKKSNSRLAGFERGK---KATV 345
DB 311 DTDQSKIFPPA-TVFFFSIEKHLERTTYEVLTDMLKEDSKFGSFRSFLTNPNATV 369
QY 346 YLIGKDFVPIADTSLGNDKILATEAISEAKKEFEKTKTIPAEVRKTLIEIPM--PD 403
DB 370 YKGISDFVGVSNSTVADADKVAQEPFLNBTADPFKQIQANPI-NYKSVLGITFMLIND 428
QY 404 KQPDQKQESLDELITDIN 421
DB 429 NDAKDNKALNELIKIN 446

```

RESULT 6
AAO15867
ID AAO15867 standard; protein; 457 AA.
XX
XX AAO15867;
AC
XX
DT 16-JAN-2003 (first entry)
XX
XX Mycoplasma hyorhinis Ag234-5 protein.
DE
XX
XX Apoptocin antigen; enzootic mycoplasma pneumoniae; mhp3; Ag234-5;
XX vaccine; Mycoplasma hyopneumoniae infection.
XX
XX Mycoplasma hyorhinis.
OS
XX
XX EPI245677-A1.
XX
XX 02-OCT-2002.
XX
XX 30-MAR-2001; 2001EP-00303030.
XX
XX 30-MAR-2001; 2001EP-00303030.
XX
XX (PF12 ) PFIZER PROD INC.
XX
XX Kling KW, Madura RA, Rosey EL;
XX
XX WPI; 2002-742716/81.
XX
XX Novel aporetocin antigens encoded by Mycoplasma hyopneumoniae for use in
XX vaccines to prevent and treat diseases caused by infection with
XX Mycoplasma hyopneumoniae in animals, especially pigs.
XX
XX Example; Fig 1; 38pp; English.
XX
XX The invention comprises the amino acid and coding sequences of Mycoplasma
XX hyopneumoniae mhp3 proteins, the invention also comprises novel
XX aporetocin antigens encoded by the M. hyopneumoniae mhp3 gene. M.
XX hyopneumoniae is a bacterial pathogen that causes enzootic mycoplasma
XX pneumonia in pigs. The mhp3 genes, proteins and aporetocin antigens of
XX the invention are useful in the manufacture of a vaccine for treating/
XX preventing a disease or disorder caused by infection with M.
XX hyopneumoniae in an animal, especially a pig. The present amino acid
XX sequence represents the Mycoplasma hyorhinis Ag234-5 protein which is
XX shown in a figure of the invention
XX
XX Sequence 457 AA;
SQ
Query Match 32.2%; Score 693.5; DB 5; Length 457;
Best Local Similarity 39.0%; Pred. No. 8.5e-47;
Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;
QY 5 ETTKEEKSAADNOKQITDVSKISGLVNERKSEIMAAKADANKHGLNMAIVTAGTVNDN 64
DB 28 ETDKEGKII-----RIPD-----NSFVKDRQAEIEKA-----KNPDFVTVLLTAGTVQDK 73
QY 65 SFNOSGWEAI-----QQLGALTG-GEITSVDSSTAELEGKYSLSLANTKNVWVLSGFGHD 119
DB 74 SFNOSIWEAVELEHVDIEKTTNLDVSOETNOSLELGKYNFANGKNKWWIILGFOGQ 133
QY 120 AFRMLKIPEN-----QQLTEKNIILGIDW---TDENVIPGRYINLTYTEAGWL 171
DB 134 EPRFKLKOTSDNGSKYDLSAEKVIIVAVDMDSKDKDLIKGHIISLYKTEEGFI 193
QY 172 AGYANASFLAKKPSDPTKSAIVGGISPAVTDFTAGIAGIKAVNLKNSDKTKYITT 231
DB 194 AGYASSKFLAYKFPNDKAKRTIAPFGGSHGAGVDFIAGFIAGIAYKNNNDPTAKVITSD 253
QY 232 DKLEINLGPVODSTERLEQLASXDKPSTLLAVASPLTEIFSDII--ANONRYILGV 289
DB 254 NNINIDTGF-ISNDKTIPTFNGIYNKS--SLVLEVASLTSVVDALKSKSKDKTKYILGV 310
QY 290 DTDOSLVTYTKNKFSTSLKNLGYSVFVSLDLVYTKKSNRLAGFEFGK---KSATV 345

```

```

DB 311 DTDOSKIFSPA-TVFFTSIEKHIGRTYQVLTIDIMLKEDSKFSGSPRSFKLTNPANATV 369
QY 346 YLGIKDRFVNIADPTSLEGNDKLTATEISAEKKEPEKRTIPEAEVKTLEIPEM---P 402
DB 370 YKGISDPEFVSNSTVADADKVRKQEFLEATADPFKKQIOANPT-NYKSVLGITPMLIND 428
QY 403 DKQPDKQOESL 413
DB 429 NDAKDNEKASL 439
RESULT 7
AAR67582
ID AAR67582 standard; protein; 432 AA.
XX
XX AAR67582;
AC
XX
XX 06-OCT-1995 (first entry)
XX
XX Cancer metastasis protein.
DE
XX
XX Cancer metastasis; clinical tests; detection.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX FT Misc-difference 55 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 99 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 140 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 140 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 319 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 420 /note= "corresponding codon TGA"
XX
XX FT Misc-difference 424 /note= "corresponding codon TGA"
XX
XX FT JF06319559-A.
XX
XX EN 22-NOV-1994.
XX
XX PF 12-MAY-1993; 93JP-00144165.
XX
XX PR 12-MAY-1993; 93JP-00144165.
XX
XX PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX WPI; 1995-040317/06.
XX
XX DR N-PSDB; AAO79124.
XX
XX DNA encoding polypeptide involved in cancer metastasis - useful for the
XX study of the mechanism of metastasis and in clinical tests.
XX
XX Claim 1; Page 2; 13pp; Japanese.
XX
XX CC AAO79124 encodes AAR67582 a protein involved in cancer metastasis, which
XX may be used for the study of metastatic mechanisms, and for clinical tests
XX to determine the presence or absence of cancer metastasis
XX
XX Sequence 432 AA;
SQ
Query Match 32.1%; Score 692.5; DB 2; Length 432;
Best Local Similarity 39.0%; Pred. No. 9.4e-47;
Matches 168; Conservative 81; Mismatches 141; Indels 41; Gaps 14;
QY 5 ETTKEEKSAADNOKQITDVSKISGLVNERKSEIMAAKADANKHGLNMAIVTAGTVNDN 64
DB 3 ETDKEGKII-----RIPD-----NSFVKDRQAEIEKA-----KNPDFVTVLLTAGTVQDK 48
QY 65 SFNOSGWEAI-----QQLGALTG-GEITSVDSSTAELEGKYSLSLANTKNVWVLSGFGHD 119

```

```
Db 49 SFNQSWEAVLEHYDQIEKTNLDVSOETNNOSELIGKYKNFLGNKNWILTLGFQOQ 108
Qy 120 AFTWMLKIPEN-----KQLFTEKNIIILGIDW--TDTEENVIPTRGVINTLYKTEBAGWL 171
Db 109 EFPKFLKQTDNSNGKYSDLAEKKVILVAVDMDLSKEDKDLKAGHPSILLYKTEBAGFI 168
Qy 172 AGVANASFLAKKPPSPDTPKRSALVIGGISPAYTDPIAGIAGIKAMNLKNSDKTKITTT 231
Db 169 AGYASSKFLAYKFPNDEAKRTIAPFGGHHGAGVTDPIAGIAGIKAMNLNDPTEKVTISD 228
Qy 232 DKIEINLGPVQDOSTYKERLEQIASKDKPSTLLAVAGLEIIFSDII--ANONRYILGV 289
Db 229 NNINIDTGF-LSNKTATFINGIVNKS--SLVLPVAGSLTSSVVDALIKSKDKTKYLLGV 285
Qy 290 DTDOSLVTYTKNKKFFTSILKNLGYSVFSVLSDLTYTKSNSRNLAGFEFGK---KSATV 345
Db 286 DTDOSKIFPPA-TYFFTSIEHGLRTIYQVLTIDWLKEDSKFLGSPRSFKLTGPAANATV 344
Qy 346 YLGIDKRFVDIADTSLGNDKKLATEAISAKKEFEEXTKTIPEAEVARKTLEIPDM--P 402
Db 345 YKGISDFFGVNSTVADADKVAQOEFLNEATADFCKQIQANPT-NYKSVLGIFPMILND 403
Qy 403 DKQPDKQOESL 413
Db 404 NDARKNEKASL 414
```

RESULT 8

AAP93343 ID AAP93343 standard; protein; 448 AA.

AAP93343;

25-MAR-2003 (revised)
01-AUG-1990 (first entry)Gene encoding the 43 kd regression-associated antigen (RA Ag) of
Mycoplasma hyorhinis.Tumour regression-associated antigens (RA Ag): in vivo imaging;
therapy monitoring; cancer therapy; Mycoplasma hyorhinis.

Mycoplasma hyorhinis.

Key Location/Qualifiers
FT Peptide 1..25
FT Protein 26..448

EPJ08265-A.

22-MAR-1989.

16-SEP-1988; 88EP-00308625.

16-SEP-1987; 87US-00097910.

11-DEC-1987; 87US-00131815.

(INGE-) INT GENETIC ENG INC.

(INGE-) INGENE INT GENETIC.

(INGE-) INGENE INT GENETIC.

Fareed GC, Sen A, Ghoshdastgi P, Li A, Lee JH;

WPI; 1989-087638/12.

N-PSDB; AAN90684.

Tumour regression-associated antigens and antibodies - used in diagnostic
tests, monitoring course of therapy and for therapy in cancer patients.

Disclosure; Page 7; 56pp; English.

Regression-associated antigens may be used in diagnostic tests, eg in

CC vivo imaging, for monitoring the course of therapy or for therapeutic
CC purposes, eg active immunisation protocols in cancer patients or drug
CC delivery systems by binding the drug to monoclonal or monospecific
CC polyclonal Ab showing specific immunoreactivity with the Ag. (Updated on
CC 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA
CC field.)

Sequence 448 AA;

Query Match 29.8%; Score 642.5; DB 1; Length 448;
Best Local Similarity 38.0%; Pred. No. 1e-42;
Matches 166; Conservative 83; Mismatches 145; Indels 43; Gaps 18;

5 ETTKEKSNQNKQITDVSIGSLVNERSEIIMAAKADANKHGLMALVTAGTINDN 64

28 ETDDEGKII-----RIPD-----NSFVDRQAEIEKA-----KMFDFNTVLLTAGTVQDK 73

65 SFNQSWEAVLEHYDQIEKTNLDVSOETNNOSELIGKYKNFLGNKNWILTLGFQOQ 119

74 SFNQSWEAVLEHYDQIEKTNLDVSOETNNOSELIGKYKNFLGNKNWILTLGFQOQ 133

120 AFTWMLKIPEN-----KQLFTEKNIIILGIDW--TDTEENVIPTRGVINTLYKTEBAGWL 171

134 EFPKFLKQTDNSNGKYSDLAEKKVILVAVDMDLSKEDKDLKAGHPSILLYKTEBAGFI 193

172 AGVANASFLAKKPPSPDTPKRSALVIGGISPAYTDPIAGIAGIKAMNLKNSDKTKITTT 231

194 AGYASSKFLAYKFPNDEAKRTIAPFGGHHGAGVTDPIAGIAGIKAMNLNDPTEKVTISD 253

232 DKIEINLGPVQDOSTYKERLEQIASKDKPSTLLAVAGLEIIFSDII--ANONRYILGV 289

254 NNINIDTGF-LSNKTATFINGIVNKS--SLVLPVAGSLTSSVVDALIKSKDKTKYLLGV 310

290 DTDOSLVTYTKNKKFFTSILKNLGYSVFSVLSDLTYTKSNSRNLAGFEFGK---KSATV 345

311 DTDOSKIFPPA-TYFFTSIEHGLRTIYQVLTIDWLKEDSKFLGSPRSFKLTGPAANATV 367

346 YLGIDKRFVDIADTSLGNDKKLATEAISAKKEFEEXTKTIPEAEVARKTLEIPDM-PDK 404

368 YKGISD-FLGVNSTVADADKVAQOEFLN-ATADFCKQIQANPT-NYKSVLGIFPMILIND 424

405 QPDQKQESLDTIDIN 421

425 DAKNERALNELIKIN 441

RESULT 9

AAY05332 ID AAY05332 standard; peptide; 428 AA.

AAY05332;

25-JUN-1999 (first entry)

Inflammatory cytokine inducer.

Inflammatory cytokine inducer; thrombocytopaenia; therapy.

Synthetic.

JPI1089582-A.

06-APR-1999.

19-SEP-1997; 97JP-00273649.

19-SEP-1997; 97JP-00273649.

(CHUS) CHUGAI PHARM CO LTD.

WPI; 1999-281057/24.

N-PSDB; AAX33847.

PT New inflammatory cytokine inducer gene and polypeptide - useful for
PT treatment of thrombocytopenia.

PS Claim 3, Page 16-17, 22pp; Japanese.

XX This sequence represents the inflammatory cytokine of the invention. The
CC inflammatory cytokine can be used in a drug, which is useful for the
CC treatment of thrombocytopenia

XX Sequence 428 AA:

Query Match 17.6%; Score 379.5; DB 2; Length 428;
Best Local Similarity 26.7%; Pred. No. 1,2e-21;

Matches 115; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

QY 31 NERKSEIMAKADANKH-----FGLMAIVTAGCTVNDSPNOSGWEAI 74
DB 27 NNDSESNISFKKIDSKTTTNNANGQVYKNAELLKLPVLTDEGKIDDSFNOSAFEL 86
QY 75 QOLGALTGETSVDSSTAELEGKYSLSANTKNVWVLSGFQGDFTFRLKIPENKOLF 134
DB 87 KAIKQGTGIEINNVES--SNFESAYNSALSAGHKIWLNGFKHQOSIKQY--IDAHEEL 143
QY 135 TEKNIILIGIDWTDENVIPTRGRYINLTYTEAGWLAGYANASFLAKGPPSDPTKSAI 194
DB 144 ERNOIKIIGIDF-DIETEVKW--FYSLOFNKESAFITGVAIASWLSAQ--DESKVVA 197
QY 195 VIGGGSIPATVDFIAGIAGIKAMNLKNSDKTKI--TTDKIEINLGFVDVDTSTKLEAQ 253
DB 198 SFGGGAFFGVTTTFEGAKGLIYVNOCH--KSKITHTSPVKLDSGF--TAGKMMT 250
QY 254 IAS-----KDKPSTLLAVAGPLTEIFPSDIIANONDRYLIGVTDGSLVYTKRNK 303
DB 251 VINNVLSSTPADVYKYNHVLISVAGPAT--FETVRLANKQGYVIGVSDQGM--QDKDR 306
QY 304 FFSILKNLGSVPSVLSL-----YTKKSNRNLAGFEGKKSATVYIGIDRF 353
DB 307 ILTSVLKHIOAVYETLLDILEKEGKPYVVDKKADKWSHFQO-----KXW 358
QY 354 VDIADTSLGNDDKL-ATEAISEAKKEFEKTKTIPAEVRKTLIEIPMP-KQPDQOE 411
DB 359 IGVAENHFSNTEBQAKINNKIKAIKMFKE---LPEDFVKYINSKALDKGNKIDVSE 414
QY 412 SLDKLTIDIN 421
DB 415 RLEHIIISAIN 424

RESULT 10

ABP71709 ID ABP71709 standard; protein; 428 AA.

XX AC ABP71709;

DT 04-APR-2003 (first entry)

XX M161 antigen amino acid sequence.

XX M161; antigen; cytotoxic T-lymphocyte inducer; CTL; cytostatic; cancer;

XX anticancer agent.

XX Mycoplasma fermentans.

OS Key Location/Qualifiers

FT M161 difference 121 /note= "Encoded by TGA; unexplained stop codon in coding

FT M161 difference 163 sequence"

FT M161 difference 185 /note= "Encoded by TGA; unexplained stop codon in coding

FT M161 difference 185 /note= "Encoded by TGA; unexplained stop codon in coding

FT Misc-difference 348 /note= "Encoded by TGA; unexplained stop codon in coding

FT M161 difference 358 sequence"

XX WO2002102402-A1.

XX 27-DEC-2002.

XX 13-JUN-2002; 2002WO-JP005916.

XX 15-JUN-2001; 2001JP-00182250.

XX (TAKE) TAKEDA CHEM IND LTD.

XX WPI; 2003-167451/16.

XX N-PSDB; AB259715.

XX M161 antigen, its peptide fragment or their salt with e.g. superior CTL-

XX inductive effect, applicable in anticancer agents combined with cancer

XX antigen for preventing or treating cancer.

XX Disclosure; Fig 1, 94pp; Japanese.

XX The invention relates to cytotoxic T-lymphocyte (CTL)-inducers containing

XX the M161 antigen, its partial peptide or their salt. The activity of CTL-

XX inducers of the invention may be described as cytostatic. The antigen or

XX its peptide fragment are applicable in anticancer agents for the

XX prevention or treatment of cancer. The antigen and its derivative have

XX superior CTL-inductive effect, giving anticancer agents that have low

XX toxicity. The current sequence represents the M161 antigen amino acid

XX sequence

XX Sequence 428 AA:

Query Match 17.6%; Score 379.5; DB 6; Length 428;
Best Local Similarity 26.7%; Pred. No. 1,2e-21;

Matches 115; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

QY 31 NERKSEIMAKADANKH-----FGLMAIVTAGCTVNDSPNOSGWEAI 74
DB 27 NNDSESNISFKKIDSKTTTNNANGQVYKNAELLKLPVLTDEGKIDDSFNOSAFEL 86
QY 75 QOLGALTGETSVDSSTAELEGKYSLSANTKNVWVLSGFQGDFTFRLKIPENKOLF 134
DB 87 KAIKQGTGIEINNVES--SNFESAYNSALSAGHKIWLNGFKHQOSIKQY--IDAHEEL 143
QY 135 TEKNIILIGIDWTDENVIPTRGRYINLTYTEAGWLAGYANASFLAKGPPSDPTKSAI 194
DB 144 ERNOIKIIGIDF-DIETEVKW--FYSLOFNKESAFITGVAIASWLSAQ--DESKVVA 197
QY 195 VIGGGSIPATVDFIAGIAGIKAMNLKNSDKTKI--TTDKIEINLGFVDVDTSTKLEAQ 253
DB 198 SFGGGAFFGVTTTFEGAKGLIYVNOCH--KSKITHTSPVKLDSGF--TAGKMMT 250
QY 254 IAS-----KDKPSTLLAVAGPLTEIFPSDIIANONDRYLIGVTDGSLVYTKRNK 303
DB 251 VINNVLSSTPADVYKYNHVLISVAGPAT--FETVRLANKQGYVIGVSDQGM--QDKDR 306
QY 304 FFSILKNLGSVPSVLSL-----YTKKSNRNLAGFEGKKSATVYIGIDRF 353
DB 307 ILTSVLKHIOAVYETLLDILEKEGKPYVVDKKADKWSHFQO-----KXW 358
QY 354 VDIADTSLGNDDKL-ATEAISEAKKEFEKTKTIPAEVRKTLIEIPMP-KQPDQOE 411
DB 359 IGVAENHFSNTEBQAKINNKIKAIKMFKE---LPEDFVKYINSKALDKGNKIDVSE 414
QY 412 SLDKLTIDIN 421

[illegible]

Qy		354	VDIADTSLSGNDKLI-ATPEALSEAKGEEFEKEKTTPAEYEVKTLLEPKMPD-KQPDKQE	411
Dd		335	IIVAHENHPSNTEEQAKINNKIKELTKMFKP-----LPDEFVKYINSDKALKDGKNKIDNVSE	390
Qy		412	SDDKLITDIN	421
Dd		391	RLEAITSAIN	400
		RESULT 12		
	ID	ADK68531		
	ID	ADK68531	standard; protein; 428 AA.	
XX	AC	ADK68531;		
XX	DT	06-MAY-2004	(first entry)	
XX	DE	Mycoplasma fermentans IL-X precursor protein #1.		
XX	KM	Immunoregulatory factor; IL-X; lymphocyte proliferation.		
OS	XX	Mycoplasma fermentans.		
FH	Key	Location/Qualifiers		
FT	Peptide	1..24		
FT	Protein	/label= Signal_peptide 25..428 /note= "Mature IL-X protein"		
PN	US6506892-B1.			
PD	14-JAN-2003.			
XX	Pf	29-OCT-1998;	98US-00182625.	
XX	Pr	29-OCT-1997;	97US-0063701P.	
PA	(WEBB/) WEBB A C.			
PA	(BLAZAR) BLAZAR B A.			
PI	Webb AC, Blazar BA;			
DR	WPI: 2003-401108/38.			
N-PeDB:	ADK68530.			
PT	New isolated polynucleotide, useful for augmenting proliferation of B and T lymphocytes.			
PS	Claim 25; SEQ ID NO 9; 42pp; English.			
CC	The invention provides novel immunoregulatory factor designated IL-X which has been isolated Mycoplasma and polnuncleotides encoding such polypeptides. The polynucleotide is useful for augmenting proliferation of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-X precursor protein.			
XX	Sequence 428 AA;			
Query Match	17.6%; Score 378.5; DB 7; Length 428;			
Best Local Similarity	26.7%; Pred. No. 1.4e-21;			
Matches 115; Conservative	88; Mismatches 156; Indels 71; Gaps 17			
Qy	31 NERKSEIMAAKADANKH-----FGLMAIYTAGTVVDNSNGSWEAI	74		
Dd	27 NNDSNNISFFKRKDISKYTTTNANGQQVNAELLKPLKITDEEGKIDKSFNOSAFAFL	86		
Qy	75 OOLGALTGCGETSVDSSTALEGGKYSSLANTNKONVVLSGFQHGDAFTRMLKI PENKQLP	134		
Dd	87 KAIINKQTGIETINNVEPS-SNFESAVNSALLSAGHKIWLVNGFGHQSIKOY--IDAHRDEL	143		
Qy	135 TEKNIILIGIDMTDETENVIPTRGYINLTYKTBEGAWLAGVANASFLAKKPSPDPYGRSAI	194		

Db 144 ERNOIKIGIDF-DIETEKYM--FYSLOFNKESAFITGVAIASWLSQ---DESKRVA 197
 QY 195 VIGGGISPAVTDPLAGLAGIKAMNLSKSDKTKI--TTDKIEINLGFVDQSTKERLEQ 253
 Db 198 SFGGGAFFGVITTFEGRAKGLIYNOQH--KSKITIHTSPVKLDSGF-----TAGEMANT 250
 QY 254 IAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLLIGVDTDSLVTYTKTKNK 303
 Db 251 VINNVLSSTPADVKYNPHVILSVAGPAT--FETVRLANKQGYIVGVDSQGMV--QDKDR 306
 QY 304 FFFSILKNLGYSPFSVLSDL-----YTKKSNRNLAGEFFGKKSATVYIGIDRF 353
 Db 307 ILTSVLGHIRKQAVETLIDLILEKEEGKPYVVDKKADKKWSHFQO-----REKW 358
 QY 354 VDIADTSLSEGNDDKL-ATEAISEAKKEFEKTKTIPAEVYRKTLEIPEMPDKOPDKQOE 411
 Db 359 IGVAENHFSTBERQAKINNKIKKAIKMFKE---LPEDFVKYINSDALKDGNKIDIVSE 414
 QY 412 SLDKLITDIN 421
 Db 415 RLEALISAIN 424

RESULT 13

ADK68537 ID ADK68537 standard; protein; 404 AA.

AC ADK68537;

XX 06-MAY-2004 (first entry)

XX Mycoplasma fermentans IL-X mature protein #2.

XX Immunoregulatory factor; IL-X; lymphocyte proliferation.

XX Mycoplasma fermentans.

XX US6506892-B1.

XX 14-JAN-2003.

XX PF 29-OCT-1998; 98US-00182625.

XX PR 29-OCT-1997; 97US-0063701P.

XX PA (WEBB/) WEBB A C.

XX PA (BLAZ/) BLAZAR B A.

XX PI Webb AC, Blazar BA;

XX WPI; 2003-401108/38.

XX N-PSDB; ADK68535.

XX New isolated polynucleotide, useful for augmenting proliferation of B and T lymphocytes.

XX Claim 13; SEQ ID NO 15; 42pp; English.

XX The invention provides novel immunoregulatory factor designated IL-X

XX CC which has been isolated Mycoplasma and polynucleotides encoding such

XX CC polypeptides. The polynucleotide is useful for augmenting proliferation

XX CC of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-

XX CC X mature protein.

XX Sequence 404 AA;

QY Query Match 17.4%; Score 375.5; DB 7; Length 404;
 Best Local Similarity 26.7%; Pred. No. 2.3e-21;
 Matches 115; Conservative 87; Mismatches 157; Indels 71; Gaps 17;

Db 31 NERKSEITMAKADANKI-----FGLMAITTAGCTVNDNSFNQSGWEAI 74
 3 NDNSENISFKEKIDISKYTTTNGKQVKNALKLKPLITIDEGKIDDKSFNQSAPFAL 62

QY 75 QOLGALTGEITSDVSTAELEKYSGLANTKNVWVLSGFQHGDAFTRLKIPENKOLF 134
 Db 63 KAIKNGTGIENNVEPS-SNFESAYNSALSAQHKIWLNGFKHQOSIKQY--IDAYNEEL 119
 QY 135 TENKIITLIGDWTDENVTGRIYINLTYTEBAGLAGAANASFLAKKPPSPDTKSAI 194
 Db 120 ERNOIKIGIDF-DIETEKYM--FYSLOFNKESAFITGVAIASWLSQ---DESKRVA 173
 QY 195 VIGGGISPAVTDPLAGLAGIKAMNLSKSDKTKI--TTDKIEINLGFVDQSTKERLEQ 253
 Db 174 SFGGGAFFGVITTFEGRAKGLIYNOQH--KSKITIHTSPVKLDSGF-----TAGEMANT 226
 QY 254 IAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLLIGVDTDSLVTYTKTKNK 303
 Db 227 VINNVLSSTPADVKYNPHVILSVAGPAT--FETVRLANKQGYIVGVDSQGMV--QDKDR 282
 QY 304 FFFSILKNLGYSPFSVLSDL-----YTKKSNRNLAGEFFGKKSATVYIGIDRF 353
 Db 283 ILTSVLGHIRKQAVETLIDLILEKEEGKPYVVDKKADKKWSHFQO-----REKW 334
 QY 354 VDIADTSLSEGNDDKL-ATEAISEAKKEFEKTKTIPAEVYRKTLEIPEMPDKOPDKQOE 411
 Db 335 IGVAENHFSTBERQAKINNKIKKAIKMFKE---LPEDFVKYINSDALKDGNKIDIVSE 390
 QY 412 SLDKLITDIN 421
 Db 391 RLEALISAIN 400

RESULT 14

ADK68538 ID ADK68538 standard; protein; 428 AA.

AC ADK68538;

XX 06-MAY-2004 (first entry)

XX Mycoplasma fermentans IL-X precursor protein #2.

XX Immunoregulatory factor; IL-X; lymphocyte proliferation.

XX Mycoplasma fermentans.

XX Key Location/Qualifiers

XX Peptide 1..24

XX Protein /label= Signal_peptide 25..428

XX /note= "Mature IL-X protein"

XX US6506892-B1.

XX 14-JAN-2003.

XX PF 29-OCT-1998; 98US-00182625.

XX PR 29-OCT-1997; 97US-0063701P.

XX PA (WEBB/) WEBB A C.

XX PA (BLAZ/) BLAZAR B A.

XX PI Webb AC, Blazar BA;

XX WPI; 2003-401108/38.

XX N-PSDB; ADK68535.

XX New isolated polynucleotide, useful for augmenting proliferation of B and T lymphocytes.

XX Claim 26; SEQ ID NO 16; 42pp; English.
 XX The invention provides novel immunoregulatory factor designated IL-X
 CC which has been isolated Mycoplasma and polynucleotides encoding such

CC polypeptides. The polynucleotide is useful for augmenting proliferation
 CC of B and T lymphocytes. The present sequence is Mycoplasma fermentans IL-
 CC X precursor protein.

XX Sequence 428 AA;

Query Match 17.4%; Score 375.5; DB 7; Length 428;
 Best Local Similarity 26.7%; Pred. No. 2.5e-21;
 Matches 115; Conservative 87; Mismatches 157; Indels 71; Gaps 17;

QY 31 NERSEIMAAKADANKH-----FGLMAIYTAGTVNDNSFNQGEAI 74
 DB 27 NNDNSNISPEKDKISKYTTTANAGKQVNAELKLPILITDGGKIDDKSFNQSAPFAL 86
 QY 75 QQLGALTGGEITVSSTAELEGKYSLSANTNKVWVLSGFQHDATFRMLKIPENKOLF 134
 DB 87 KAIKQKGIEINNVEPS-SNFSAYNSALSGHKIWLNGFKHQOSIKQY--IDAYREEL 143
 QY 135 TEKRIIILGIDMTENVIPTRGYINLTYTEEGWLAGYANASFLAKKPPSDPTKSAI 194
 DB 144 ERNOKIKIGIDF-DIETREYKW--FYSLOFNKESAFITGVALASWLSQ--DESKRVVA 197
 QY 195 VIGGISPATVDTFAGYLAGIKAWNLKNSDKTKI--TTDKIENLGPVDSTKRELEQ 253
 DB 198 SFGGGAFFGVTTNEGFAKGLIYNQKH--KSKIIYHTSPVKLDSGF-----TAGEKANT 250
 QY 254 IAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVTDOSLVTYTKTNK 303
 DB 251 VINNVLSSTPADVKNPHVILSVAGPAT--PEYVRLANKGGYVIGVSDQGM--QDKDR 306
 QY 304 FFFSILKMLGYSVPSVSLD-----YTKKSNRNLAGEFGKKSATVYLGIDRF 353
 DB 307 ILTSVLKRIKQAVYETLIDLLEKEEGKPYVVDKKADKKMSHFGTQ-----KEKW 358
 QY 354 VIDIDTSLGNDKTL-ATEAISEAKKEFEKTKTTPAEVVKTELEIPMPD-KQDPKQOE 411
 DB 359 IGVENHPSNTEBOAKINNKIKELKMFKE---LPEDFVYINSDKALKGKNKIDNVSE 414
 QY 412 SLDKLITDIN 421
 DB 415 RLEALISAIN 424

RESULT 15
 AAM22727 standard; protein; 429 AA.

XX AAM22727;

XX AC 26-FEB-1998 (first entry)

XX DE Membrane protein M161Ag.

XX KM M161Ag; membrane protein; leukaemia; treatment.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 121 /note= "selenocysteine"

FT Modified-site 163 /note= "selenocysteine"

FT Modified-site 185 /note= "selenocysteine"

FT Modified-site 348 /note= "selenocysteine"

FT Modified-site 358 /note= "selenocysteine"

XX JF09157295-A.

XX PD 17-JUN-1997.

PF 05-DEC-1995; 95JP-00344504.

XX 05-DEC-1995; 95JP-00344504.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 1997-369470/34.

XX N-PSDB; AATV5133.

XX Membrane protein M161Ag - useful in development of leukaemia treating agent.

PS Claim 1; Fig 1; 6pp; Japanese.

CC This sequence is a membrane protein designated M161Ag. Oligonucleotide
 CC probes for the cDNA were designed from putative N-terminal M161Ag peptide
 CC sequences isolated from P39 cells. M161Ag can be used in the development
 CC of a leukaemia treating agent

XX Sequence 429 AA;

Query Match 17.3%; Score 372; DB 2; Length 429;
 Best Local Similarity 26.8%; Pred. No. 4.8e-21;
 Matches 117; Conservative 88; Mismatches 167; Indels 64; Gaps 18;

QY 13 ADNQNKOIT---DVSKISGLVNERKSEIMAAKADANKHFGMAIYTAGTVNDNSFNQ 68
 DB 26 ANNDNSNISPEKDKISKYTTTANAGKQVNAEL-----LKLKPVLLITDGGKIDDKSFNQ 80
 QY 69 SCWEATIQQLGALTGGEITVSSTAELEGKYSLSANTNKVWVLSGFQHDATFRMLKIP 128
 DB 81 SAFALKAIKQKGIEINNVEPS-SNFSAYNSALSGHKIWLNGFKHQOSIKQY--ID 137
 QY 129 ENKQLFTEKRIIILGIDMTENVIPTRGYINLTYTEEGWLAGYANASFLAKKPPSDP 188
 DB 138 AHREBELERNOIKIIGIDF-DIETREYKW--FYSLOFNKESAFITGVALASWLSQ--DE 191
 QY 189 TKGSAIVIGGISPATVDTFAGYLAGIKAWNLKNSDKTKI--TTDKIENLGPVDST 247
 DB 192 SKRVVAFSGGAFPGVTTNEGFAKGLIYNQKH--KSKIIYHTSPVKLDSGF-----TA 244
 QY 248 KERLEQIAS-----KDKPSTLLAVAGPLTEIFSDIIANQNDRYLIGVTDOSL 297
 DB 245 GERKNIVINNVLSTPADVKNPHVILSVAGPAT--PEYVRLANKGGYVIGVSDQGM-- 301
 QY 298 TKTNKEFFSILKMLGYSVPSVSLD-----YTKKSNRNLAGEFGKKSATVYL 347
 DB 302 -QDKDRILTSVLKRIKQAVYETLIDLLEKEEGKPYVVDKKADKKXSHFGTQ----- 354
 QY 348 GIKDRFYDIADTSLGNDKTL-ATEAISEAKKEFEKTKTTPAEVVKTELEIPMPD-KQ 405
 DB 355 --KEKXIGVABNQFSNTEBOAKINNKIKELKMFKE---LPEDFVYINSDKALKGKNK 408
 QY 406 PDKQOESLDKLTIDIN 421
 DB 409 IDNVSERLAALISAIN 424

Search completed: December 18, 2004, 01:22:25
 Job time : 112.476 secs


```

; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...374
; SEQUENCE DESCRIPTION: SEQ ID NO: 5084
US-09-107-532A-5084

```

Query Match	9.4%	Score 202	DB 4	Length 374
Best Local Similarity	25.2%	Pred. No. 1.6e-10		
Matches 95	Conservative 63	Mismatches 133	Indels 86	Gaps 19

[illegible]

RESULT 9

```

US-09-071-035-80
; Sequence 80, Application US/09071035
; Patent No. 648043
;
GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
;
CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
;
ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
INFORMATION FOR SEQ ID NO: 80:

```

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 339 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;
US-09-071-035-80

```

Query Match	8.3%	Score 178	DB 4	Length 339
Best Local Similarity	25.1%	Pred. No. 2.56	08	
Match 103, Conservative	51	Mismatches 140	Indels 116	Gaps 20

```

Qy      33 RKEIIMAAARADANKRHFJGIMMAIVTQAGVWNNSFNQSGWEAIQOLGA---LTGGE-----84
Db      7 KTAESGGGAGDA-----AHSAVLIITDGTGVDDKSFNQSGWBELOQMGKEHDLPGSGKGYA 62

Qy      85 ITSVDSS--TAELEGYSGLANTKNKWNVLVSGFOHGDALFRWLKIPENKOLFEYKNIIL 142
Db      63 IGSNDADADYTTNIDQAVSSKEFTTIFGI---GYLLKDA-----ISSAADANPDPTNVL 112

Qy      143 GIDMTDTEWVITGKRYINLTYKTEBAGLAGYANASFLAKKFPSSDDPKRSAYIVGGISIP 202
Db      113 DDQIDGRKNV-----VSATFEDNEAAALAGVAAA-----NETKTNKGVFGVGGEEG 157

Qy      203 AVTD-PIAGYLAGIKAMWLKNSDKKTKITTYDKIEINLGFDVQDSTYERLEQJASKDKPS 261
Db      158 VVIDRFQAGFEKGV-ADAKELGKEITYDT-----KYAASFDPDA 196

Qy      262 TLLAVAGPLTEIFSDI-----ANONDR-YLIGVDTQSL--VYT 298
Db      197 KGRLLAAMVYQNGVDIIFHAGSAGTQGVFOEAKDLNBSGSGDKVWIVGDRDDQADQKY- 255

Qy      299 KTK-----NKFTSLIKNLGYSVSVSLDYTKKSNSBNLAGFPFGKKSATVYLGIDRF 353
Db      256 KTKDQKEDPFTLTSTLKGVGTAVODIARALEDK-----FPGEHLYV-GLKDG 304

Qy      354 VDIADTSLSGNDKCLATEAISAKKEFEKTKTIPAEVTKTLEIIPMPD 403
Db      305 VDLTLDYLL--NDKT-----KEAVKTAQDKVISGDVAVPEKPE 339

```

RESULT 10

US-08-396-957A-5
Sequence 5, Application US/08396957A
Patent No. 5780041
GENERAL INFORMATION:
APPLICANT: SIMPSON, WARREN, SCHWAN, TOM G.
TITLE OF INVENTION: ANTIGENIC PROTEINS AND
TITLE OF INVENTION: GENES ENCODING SAME OF BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,957A
FILING DATE: 01-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/020,245
FILING DATE: 19-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/664,731
FILING DATE: 05-MAY-1991

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/487,716
; FILING DATE: 05-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4018US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 341
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Borrelia burgdorferi
; STRAIN: Sh-2-82
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: p39a
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: p39a protein sequence
;
US-08-396-957A-5

Query Match      8.3%; Score 178; DB 1; Length 341;
Best Local Similarity 26.9%; Pred. No. 2.5e-08;
Matches 101; Conservative 58; Mismatches 124; Indels 92; Gaps 25;

QY 44 ANKHFGNMAIVTGGTVNDNSFNOSGWEAIOQLGALGHEITSVDSSTALEGKYS-- 101
DB 24 SSKKIKISMV--DGVLDKSFNSANEALLRLKDPENIEEYFS--CAISGYSSYV 78
QY 102 --LANTNKN--VWVLSGFOHGAFTRWLKI PENKQLFTEKNII--ILGIDMTDENVI 153
DB 79 SDLNLKRNQSDLIW-LVGYMLTDA--SLVSSENPKI--SYGIIDPIYGDVQIPEMLI 133
QY 154 PTGRYINLTYTEAGWLAGY--ANASFLAKKFPSPDPTKSAIYIG--GGISPAVTD-FI 208
DB 134 A-----VVFVEGAGFLAGYIAAKKSPSGK-----IGPIGKMGKNIIVDAFR 174
QY 209 AGYLAGIKAWMLKNSDKTKITTDKIEINLGFVDVDTSTKERLBQIASK--DKPSTILAV 266
DB 175 YGYSAGAKYAN-KDIEIISSEYNSGFSVDIG-----RTIASKMTSGKIDIVHF 221
QY 267 AGPLTEI-FSDIIANOND-RYLIGVDTQSLVYTKNKKFPTSILKNLGVSVFSLDLY 324
DB 222 AAGLAGIGVIEAANLGDGYVVIIGADQDS--YLA PKV-FITSVIKNIGDALYILTGE-Y 277
QY 325 TKKNSRNLAFEGEGKKSATYILGIKDFVADIATSLSEGNDKALATEAISAKKEFEKKT 384
DB 278 IKNNNV-----WEGSK--VVQMGIRDGVIGLIPNAN-----EF-EYI 310
QY 385 KTI PAEVRKTLPI 399
DB 311 KVLERKIVNKEIIV 325

RESULT 11
US-09-071-035-78
; Sequence 78; Application US/09071035
; Patent No. 6448043
```

```

;
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooker
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-071-035-78

Query Match      8.3%; Score 178; DB 4; Length 361;
Best Local Similarity 25.1%; Pred. No. 2.8e-08;
Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

QY 33 RKSEIIMAKADANKHFGNMAIVTAGTVNDNSFNOSGWEAIOQLGA--LTGSE----- 84
DB 29 KTAESGGKGDGDA--AHSAVIITDTGVDKSFNOSWEGQLQWGEKHDLPBGSKGYAY 84
QY 85 ITSVDSS--TAELGKYSLSLANTNKNVWVLSGFOHGAFTRWLKI PENKQLFTEKNIIIL 142
DB 85 IQSNDADVYTNIDQAVSSKFNITFGI--GYLLKDA-----ISSAADANPDTNVL 134
QY 143 GIDWTDENVIPTGRYINLTYTEAGWLAGYANASFLAKKFPSPDPTKSAIYIGGISP 202
DB 135 DDQIDGKKNV-----VSATPFRODEAAVLAGVAAA-----NETKTKVGVGVGEBG 179
QY 203 AVTD-PIAGYIAGIKAWMLKNSDKTKITTDKIEINLGFVDVDTSTKERLBQIASKKPS 261
DB 180 VVIDRFOAGFEKGV-ADAARELKEITVDT-----KYAASFPADPA 218
QY 262 TLAAVAGPLRIEIPSDII-----ANQNDR-YLIGVDTDSL--VYT 298
DB 219 KGRKLAAMQONGVDIIFHASGATGQGVFOEAKDLNMSGGDKVWVIGVDRDQADGKY- 277
QY 299 KTK-----NKEFTSILKNLGVSVFSLDLYTKKNSRNLAFEGEGKKSATVYLGIDRF 353
DB 278 KTDGKEDNFTLSTLGVGTAVQDIANRLAEK-----FPGGEHLVY-GLMDGG 326
QY 354 VDIADTSLSEGNDKALATEAISAKKEFEKKTII PAEVRKTLPIPM 403
DB 327 VDLTDGYL--NDKT-----KEAVKTAQDKVIGDVVPEKPE 361

RESULT 12
```

US-09-134-000C-6005
 ; Sequence 6005, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6005
 ; LENGTH: 375
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000C-6005

Query Match 8.3%; Score 178; DB 4; Length 375;
 Best Local Similarity 25.1%; Pred. No. 3e-08;
 Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

QY 33 RKSEIMAKADANKHFGIANNALVAGGVNDNSFNQSGWEALIQOLGA---LTGGE----- 84
 DB 43 KTAESGGGKGA---AHSAVIITDTCGVDPKSFNQSWEGLQWKEHDLPEGSKGAY 98
 QY 85 ITVDSS--TAELEGKSSLANTKNVVLSGFQHGAFTRMLKIPENKQLFEKNIIIL 142
 DB 99 IQSNDAADYNTINDQAVSSKFNITFGI---GYLLKPA-----ISSADANPNTNVL 148
 QY 143 GIDMTDENTYIPGRVYNLTYTEEAGMLAGYANASFLAKKPSDPKRSATVIGGISP 202
 DB 149 DDQIDGKNV-----VSATFRDNEAAYLVGVA---NEFTKNVGVGSEEG 193
 QY 203 AVTD-PIAGYLAGIKAWNLKNSDKKITTTDKIEINLGFVDSTKERLEQIASKDKPS 261
 DB 194 VVIDRFGAGFEKV-ADAAMELGKETVDT-----KYAASFADPA 232
 QY 262 TLAAVAGPLTEISDIT-----ANONDR-YLIGVDTDSL--VYT 298
 DB 233 KGAALAAWYQNGVDIIFHAGATGCGVFOEAKDLNMSGSGDKWVIGVDRDQADGKY- 291
 QY 299 KTK-----NKFSTILKNLGVSVFVSDLYTKKSNRNLAGFEFGKKSATVYIGIDRF 353
 DB 292 KTKDGEDNFTLTSTLKGVTAVODIANRLDEK-----FPGGEHLVY-GLKDDG 340
 QY 354 VDIADTSLSGNDKRLATEAISEAKKEFEKTKTIPAEVKTLEIPMPD 403
 DB 341 VDLTDGL-NDKT-----KEAVKTAKDVIIGDVVPEKPE 375

RESULT 13
 ; US-09-134-000C-6004
 ; Sequence 6004, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6004
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000C-6004

Query Match 8.1%; Score 174; DB 4; Length 362;
 Best Local Similarity 24.1%; Pred. No. 6.7e-08;
 Matches 96; Conservative 67; Mismatches 132; Indels 104; Gaps 20;

QY 40 AKADAN-----KFGIANNALVAGGVNDNSFNQSGWEALIQOLGALTGEITSVDSSTAS 94
 DB 32 SKKNATKGPQH---SVWVDTAGIDDKSFNQSMEGQEWG--KEHLPFGPGQAY 86
 QY 95 LEKGYSLANTNKQWVWLSGFO--HGDAPFRMLKIPENKQLFEKNIIIGIDMTDENV 152
 DB 87 IQSNEASDYTSNIDQAISSQFKTIFGIGYLLKNAVADANPEINFLV---DITVN- 141
 QY 153 IPTGRVYNLTYTEEAGMLAGYANASFLAKKPSDPKRSATVIGGISPATDFIAGY 211
 DB 142 -GKNVNASATFRNESAIVLAGVAAA-----NTTKNVGFIGVSEPVIGRFGAGF 191
 QY 212 LAGIKAWNLKNSDKKITTTDKIEINLGFVDSTKERLEQIASDKPESTLLAVAGPLT 271
 DB 192 EKV-----ADAGK-----LGKDIQITSTVAGTFADASKGR-----ALASMY 230
 QY 272 EITSDII-----ANONDRYLIGVDTDSL--VYT-----KTGKRF 304
 DB 231 QAGADITTHAAATGCGIFOEAKALNETGSKDRVWVIGVDRDQEDGKITTKDGDNL 290
 QY 305 FTSILKNLGVSVFVSDLYTKKSNRNLAGFEFGKKSATVYIGIDRFDIADTSLGDN 364
 DB 291 LASTIGVNLAV-KKISDLALDEK-----FPGGEH---LTYGLKDDGVD----- 330
 QY 365 DKXLATEAIS-EAKKEFEKTKTIPAEVKTLEIPMPD 402
 DB 331 -LTTEALSDQAKTAVKEKEQIISGDVK---VPDQ 362

RESULT 14
 ; US-09-182-625F-6
 ; Sequence 6, Application US/09182625F
 ; Patent No. 6506892
 ; GENERAL INFORMATION:
 ; APPLICANT: Webb, Andrew C.
 ; APPLICANT: Blazar, Beverly A.
 ; TITLE OF INVENTION: Polynucleotides Encoding a Mycoplasma Protein Involved in Cell Gtc
 ; FILE REFERENCE: BLAZ-101XCI
 ; CURRENT APPLICATION NUMBER: US/09/182,625F
 ; PRIOR FILING DATE: 1998-10-29
 ; PRIOR APPLICATION NUMBER: US 60/063,701
 ; NUMBER OF SEQ ID NOS: 16
 ; SEQ ID NO 6
 ; LENGTH: 166
 ; TYPE: PRT
 ; ORGANISM: Mycoplasma fermentans
 ; FEATURE:
 ; OTHER INFORMATION: Peptide
 ; US-09-182-625F-6

Query Match 8.0%; Score 173; DB 4; Length 166;
 Best Local Similarity 27.3%; Pred. No. 2.4e-08;
 Matches 50; Conservative 37; Mismatches 56; Indels 40; Gaps 8;

QY 197 GGGISPAVTFPIAGYLAGIKAWNLKNSDKKITTTDKIEINLGFVDSTKERLEQIA 255
 DB 2 GCGAFPGVTFNFGPAKGIYVNGK--KSKLYHSPVLDGDF-----TAGKKNVTI 54
 QY 256 S-----KKESTLLAVAGPLTEISDIANONDRYLIGVDTDSLVTYTKNKKF 305
 DB 55 NNVLSTPADVKNPHYLIVSAGPAT--FEFTVLANKGQYIVDSQGMV--QDKDRIL 110
 QY 306 TSLIKNLGVSVFVSLD-----YTKKSNRNLAGFEFGKKSATVYIGIKDRFVD 355
 DB 111 TSVLKHITQAVETTLDDLLEKEBGKPYVYVKKAKDKKMSHFQTO-----KEKMG 162

QY 356 IAD 358
Db 163 VAE 165

RESULT 15

US-08-396-957A-4
Sequence 4, Application US/08396957A
Patent No. 5780041
GENERAL INFORMATION:
APPLICANT: SIMPSON, WARREN; SCHMAN, TOM G.
TITLE OF INVENTION: ANTIGENIC PROTEINS AND
TITLE OF INVENTION: GENES ENCODING SAME OF BORRELLIA BURGDOFFERI.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,957A
FILING DATE: 01-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/020,245
FILING DATE: 19-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/664,731
FILING DATE: 05-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/487,716
FILING DATE: 05-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4018US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 339
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Borrelia burgdorferi
STRAIN: Sh-2-82
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
ORGANELLE:
FEATURE:
NAME/KEY: p39'
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: p39' protein
OTHER INFORMATION: sequence.
US-08-396-957A-4

Query Match 7.5%; Score 161; DB 1; Length 339;
Best Local Similarity 22.8%; Pred. NO. 1e-06;
Matches 85; Conservative 68; Mismatches 116; Indels 84; Gaps 18;

QY 59 GTVNDNSFNOSGMEAIQQLALTCGETSVDSST---AELEGKYSLLANTNKNVWVLSSG 114
Db 37 GTFDKSFNESAALNGVAKVBEFKIELVKSSSNYSLSDEG---LKDAGSDLIWLIG 92
QY 115 FOHGD-AFTFWLKIPEKQOLFTEKNIIIGIDWTDTEENVIPTRGINIITYKTEBAGWLAG 173
Db 93 YRFEDVAKVAALQNPDMKVAI-----IDPIYSNDPIP-ANLVGMTFRAQGAFLTG 142
QY 174 YANASFLAKKPPSDPTKSAIVIGGGISPAVTD-FLAGIYLAGIKAMNLKNSDKTKTTTD 232
Db 143 YIAAKL-----SKTGKIGFLGGIEGIVDAFRYGEAGAKVYAN-----KDIKISTQ 188
QY 233 KI-----EINIGFDVQDISTERLEQLASKKPSLLAVAGPLIEIFSDIITANQ--NDRYL 286
Db 189 YIGSFADLEAGRSVATRMYSDEID-----IHHAGLGIGIGAIIEVAKELGSGHYI 238
QY 287 IGVDTQSLVYTTKKNKFTSILKNLGSVFSVLSDLTYTKKSNSRNLAGFEGKKSATVY 346
Db 239 IGVDEDDAYL---APDVIITSTKDVGRAL-----NIFT--SNHLKTNTEGCGK---LIN 285
QY 347 LGIKRFPVDIADTSLBGNDKLATETALSEAKKEPEEKTKTIPAEVRKTLIEIPMPDKOP 406
Db 286 YGLKEGVVGFV-----RNPWMISF-----ELEKEIDWLSKIKIKEIIVP----- 325
QY 407 DKQESLDKLITD 419
Db 326 -SNKESTYEKFLKE 337

Search completed: December 18, 2004, 01:28:36
Job time : 30.0389 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:26:40 ; Search time 97.7643 Seconds
(without alignments)
1548.256 Million cell updates/sec

Title: US-09-676-249D-4
Perfect score: 2155
Sequence: 1 MMDETTEKESADNQNKQI.....KQPKQESIDKLITDINNLL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: Published Applications AA:*

- 1: /cgn2_6/ptodaca/2/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodaca/2/pubppa/US09C_PUBCOMB.pep.*
- 3: /cgn2_6/ptodaca/2/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodaca/2/pubppa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodaca/2/pubppa/US07_PUBCOMB.pep.*
- 6: /cgn2_6/ptodaca/2/pubppa/US08_NEW_PUB.pep.*
- 7: /cgn2_6/ptodaca/2/pubppa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodaca/2/pubppa/US09A_PUBCOMB.pep.*
- 9: /cgn2_6/ptodaca/2/pubppa/US09B_PUBCOMB.pep.*
- 10: /cgn2_6/ptodaca/2/pubppa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodaca/2/pubppa/US09D_PUBCOMB.pep.*
- 12: /cgn2_6/ptodaca/2/pubppa/US09E_PUBCOMB.pep.*
- 13: /cgn2_6/ptodaca/2/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodaca/2/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodaca/2/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodaca/2/pubppa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodaca/2/pubppa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodaca/2/pubppa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodaca/2/pubppa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodaca/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	220.5	10.2	350	US-10-474-792-166	Sequence 166, Appl
2	216.5	10.0	330	US-10-451-337-12	Sequence 12, Appl
3	216.5	10.0	330	US-10-451-337-14	Sequence 14, Appl
4	216.5	10.0	330	US-10-451-337-39	Sequence 39, Appl
5	216.5	10.0	330	US-10-451-337-40	Sequence 40, Appl
6	216.5	10.0	330	US-10-451-337-41	Sequence 41, Appl
7	213.5	9.9	330	US-10-451-337-16	Sequence 16, Appl
8	213.5	9.9	330	US-10-451-337-42	Sequence 42, Appl
9	209	9.7	344	US-10-451-337-6	Sequence 6, Appl
10	202	9.4	328	US-09-765-272-8	Sequence 8, Appl
11	202	9.4	350	US-09-769-787-132	Sequence 132, Appl
12	178	8.3	339	US-09-071-035-80	Sequence 80, Appl
13	178	8.3	339	US-10-206-576-80	Sequence 80, Appl

14	178	8.3	361	US-09-071-035-78	Sequence 78, Appl
15	178	8.3	361	US-10-206-576-78	Sequence 78, Appl
16	174	8.1	347	US-10-451-337-8	Sequence 8, Appl
17	174	8.1	347	US-10-474-792-180	Sequence 180, Appl
18	173	8.0	166	US-09-760-541-6	Sequence 6, Appl
19	168	7.8	357	US-10-282-122A-60484	Sequence 60484, A
20	151.5	7.0	223	US-09-769-736-48	Sequence 48, Appl
21	148	6.9	797	US-10-156-761-10907	Sequence 10907, A
22	137.5	6.4	889	US-09-952-267-15	Sequence 15, Appl
23	133.5	6.2	1786	US-09-742-096-3	Sequence 3, Appl
24	133.5	6.2	1787	US-10-415-253-2	Sequence 2, Appl
25	127.5	5.9	496	US-10-451-467A-432	Sequence 432, App
26	127	5.9	2045	US-10-282-122A-74463	Sequence 74463, A
27	127	5.9	2059	US-10-474-792-62	Sequence 62, Appl
28	124.5	5.8	998	US-10-282-122A-70450	Sequence 70450, A
29	121.5	5.6	1073	US-10-193-764-45	Sequence 45, Appl
30	121.5	5.6	1079	US-10-193-764-43	Sequence 43, Appl
31	121.5	5.6	1849	US-10-637-544-2	Sequence 2, Appl
32	121	5.6	852	US-10-282-122A-62892	Sequence 62892, A
33	121	5.6	902	US-10-437-963-105564	Sequence 105564, A
34	120.5	5.6	2457	US-10-282-122A-49854	Sequence 49854, A
35	120	5.6	1946	US-10-282-122A-62947	Sequence 62947, A
36	119	5.5	2285	US-09-932-183A-2	Sequence 2, Appl
37	118.5	5.5	1963	US-10-282-122A-73978	Sequence 73978, A
38	118.5	5.5	6641	US-10-282-122A-70580	Sequence 70580, A
39	118.5	5.5	10203	US-10-661-809-23	Sequence 23, Appl
40	118	5.5	1031	US-10-282-122A-54611	Sequence 54611, A
41	117.5	5.5	392	US-10-194-489-22	Sequence 22, Appl
42	117.5	5.5	719	US-10-239-610-2	Sequence 2, Appl
43	117.5	5.5	1098	US-09-797-862-32	Sequence 32, Appl
44	116.5	5.4	571	US-10-282-122A-53248	Sequence 53248, A
45	116.5	5.4	1104	US-09-797-385-4	Sequence 4, Appl

ALIGNMENTS

US-10-474-792-166	US-10-474-792-166
Sequence 166, Application US/10474792	Sequence 166, Application US/10474792
Publication No. US20040236072A1	Publication No. US20040236072A1
GENERAL INFORMATION:	GENERAL INFORMATION:
APPLICANT: Olmsted, Stephen	APPLICANT: Olmsted, Stephen
APPLICANT: Zagursky, Robert	APPLICANT: Zagursky, Robert
APPLICANT: Nickbarg, Elliot	APPLICANT: Nickbarg, Elliot
TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES	TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
FILE REFERENCE: AM 100399	FILE REFERENCE: AM 100399
CURRENT APPLICATION NUMBER: US/10/474, 792	CURRENT APPLICATION NUMBER: US/10/474, 792
CURRENT FILING DATE: 2003-10-14	CURRENT FILING DATE: 2003-10-14
NUMBER OF SEQ ID NOS: 674	NUMBER OF SEQ ID NOS: 674
SOFTWARE: PatentIn version 3.0	SOFTWARE: PatentIn version 3.0
SEQ ID NO 166	SEQ ID NO 166
LENGTH: 350	LENGTH: 350
TYPE: PRT	TYPE: PRT
ORGANISM: Streptococcus pyogenes	ORGANISM: Streptococcus pyogenes
US-10-474-792-166	US-10-474-792-166
Query Match	10.2%; Score 220.5; DB 17; Length 350;
Best Local Similarity	27.8%; Pred. NO. 2.6e-08;
Matches 103; Conservative 49; Mismatches 150; Indels 69; Gaps 17;	Matches 103; Conservative 49; Mismatches 150; Indels 69; Gaps 17;
QY	39 AAKADAKHFGIMNAIVTAGTVNDSFNOSGMAIQOLGALTS-GEITSVDSATLEEG 97
DB	26 ASKGGASGKTDLVKAMVTDGVDKSFNSAMEGLSKEMGLQKGTGEDYQSTSES 85
QY	98 KYSLANTNNKVVWLTSGFQ--HGDAFTRMUKIPENKQLFTEKNIIILGIDMTDENVIPT 155
DB	86 EVA-----TNDLTAVNSGVQILYIGIFA--LKDAIAKAAGNNEGKVFYIID---DIEGK 135
QY	156 GRYINLTYSKEEAGMLAGYANASTLAKFPSDPKRSALV-VIGGIPATVDTFAGVLAG 214
DB	136 DNVAIVTFADHEAAYIAGIAAK-----TTKTYGVGVGMEGTVITRFEKFPBAG 186

```

QY      215 KKAUMLNKSDKKTKTITTKIEINLGE-----DVQDSTHERLEQIASKOKPSTLLAVAG 269
Db      187 VKS-----VDDTIOVKVDAGSFGDAKGTIAAOYAAGAD--VITYAAGT 232

QY      270 LTELPSDIANONDR-----YLIGVDITQ--SLVYT-----KTKNKFPTSLNLNLSYAF 317
Db      233 GAGVENEKALINERKSEADKWAVIGVDROQKBEKGTISDQGEANVVLASSIVEGRVQ 292

QY      318 SVLSDLYTKKSNRNLACGFEFGKKSATVYLGIKDFVADIADTSLBGNDKKLATEALSEAK 377
Db      293 LINKOVADKK-----FPGGK--TIVY-GLKDGVEIATLT---NVSKEAVKAIKEAK 337

QY      378 KEFEKTKTIP 388
Db      338 AKIKSGDIKVP 348

```

```

RESULT 2
US-10-451-337-12
; Sequence 12, Application US/10/451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denis
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Josee
; APPLICANT: RIOUX, Stephane
; APPLICANT: RHEULT, Patrick
; TITLE OF INVENTION: STEPHANOCCUS PYOGENES ANTIGENS AND
; TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS
; FILE REFERENCE: 12806-24PCT
; CURRENT APPLICATION NUMBER: US/10/451,337
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 330
; TYPE: PRF
; ORGANISM: S. pyogenes
; US-10-451-337-12

```

Query Match	10.0%;	Score 216.5;	DB 15;	Length 330;
Best Local Similarity	28.3%;	Pred. No. 4.7e-08;		
Matches 102;	Conservative 46;	Mismatches 143;	Indels 69;	Gaps 17;

[illegible]

```

RESULT 3
US-10-451-337-14
; Sequence 14, Application US/10451337
; Publication No.: US20040097706a1
; GENERAL INFORMATION:
; APPLICANT: SHIRE BIOCHEM INC.
; APPLICANT: MARTIN, Denis
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: HAMEL, Josee
; APPLICANT: RIOUX, Stephanie
; APPLICANT: REHAULT, Patrick
; TITLE OF INVENTION: STEPTOCOCCUS PYOGENES ANTIGENS AND
; TITLE OF INVENTION: CORRESPONDING DNA FRAGMENTS
; FILE REFERENCE: 12806-24PCT
; CURRENT APPLICATION NUMBER: US/10/451.337
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: US 60/256,940
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ. ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 330
; TYPE: PRT
; ORGANISM: S. pyogenes
US-10-451-337-14

```

Query Match	10.0%;	Score 216.5;	DB 15;	Length 330;
Best Local Similarity	28.3%;	Pred. NO. 4.7e-08;		
Matches 102;	Conservative 46;	Mismatches 143;	Indels 69;	Gaps 17;

Qy 39 AAAYADNKHFGNNMILVWAGGVNDNSNFGSGWEIIOOLGALUT-GEITSVDSSTAELEG 97
Db 19 ASHGASGSKTDLKVAAVVDVTGGVDKXSNQSGAMEGLOSWEKMGLOKGTGPEDYFQSGTSES 78
Qy 98 KYSSLANTKKNVWVLSGFO--HGDAFTFWLKI PENKOLFTEKNIILIGIDMTOTENVIPT 1555
Db 79 EYA-----THUDTAVSGGYOLAYIGIFA--LMDALAKAAGNDEGKVFYIID----DIEEG 128
Qy 156 GRYINLTYKTEAGWLAGYANASFLAKKFPSPDPTKSAI-VIGGSIAPVTDFTAGYLAG 214
Db 129 DNVA5TTFADHHEAAYLAGIAAAK-----TTKTKTVGVSGMEGTVTRFEKGEAG 179
Qy 215 IKAMNLKNSDKTKYKITTDDKIEINLGF-----DVODSTKEHLEOIASKDXPSTLLAVAGP 268
Db 180 VKS-----VDDTIGYKVVDYAGSPFDAKAGKTIAAQYAAAD--VIQAAAGT 225
Qy 270 LTFEFDIIANQNR-----YLIGVTDQ--SLVYT---KTQNKFFSTLKLGLGSVF 317
Db 226 GAGVFNEAKAINEKSEADKVVWIGVDPDQDEKRTYSKQGEANFVLA5SIKEVGAVQ 285
Qy 318 SVYLSDDLYTKKSSNRNLAGEPEGKSKSATYLLGIDKDPVADIATSLLEGNDKDLATPAISAEK 377
Db 286 LINKOVADKK-----PQSG--TTY-GIKQSGEVIATT---NVSKEAVKALIEAK 330

RESULT 4
US-10-451-337-39
Sequence 39 Application US/10451337
Publication No. US20040097706A1
GENERAL INFORMATION:
APPLICANT: SHIRE BIOCHEM INC.
APPLICANT: MARLIN, Denis
APPLICANT: BRODEUR, Bernard R.
APPLICANT: HAMEL, Josee
APPLICANT: RIOUX, Stephanie
APPLICANT: REHAULT, Patrick
TITLE OR INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
TITLE OR INVENTION: CORRESPONDING DNA FRAGMENTS
FILE REFERENCE: 12806-24PPT
CURRENT APPLICATION NUMBER: US/10/451.337
CURRENT FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: US 60/256,940
PRIOR FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 42
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 39
 LENGTH: 330
 TYPE: PRF
 ORGANISM: S. pyogenes
 US-10-451-337-39

Query Match 10.0%; Score 216.5; DB 15; Length 330;
 Best Local Similarity 28.3%; Pred. No. 4.7e-08;
 Matches 102; Conservative 46; Mismatches 143; Indels 69; Gaps 17;

QY 39 AAKADANKHFGMLNMAIYTAGTVNDNSFNOSGMEAIQOLGALTG-GEITSVDSSTAIEG 97
 DB 19 ASKGAAGSGKTDLKAVAMVTDGVDKSFNOSAMEGLSQWKEMLQKGTGDFYQSTSES 78
 QY 98 KYSLANTNKNVWVLSGFQ--HGDAFTRWLKIPENKQILFTEKNIIILGIDMTDTEVNIPT 155
 DB 79 EYA-----TNDTAVSGGYQLYIGIGA--LKDAIAKAGNBEKVFYIID----DIEGK 128
 QY 156 GRVINTLYKTEBAGMLAGVANNASFLAKKPSDPTKRSAL-VIGGISPATVDTPIAGYIAG 214
 DB 129 DNVAATVPADHEAAYLAGIAAIAK-----TTKTTGVPFGMEGTVITREKGEAG 179
 QY 215 IKAWNLKNSDKKTKITTDKIEINLGF----DVODSTKERLEQIASKDXPSTLLAVAGP 269
 DB 180 VKS-----VDDTQVAVDAGSFGDAKAGKTIAAQAAGAD--VIYQAAGGT 225
 QY 270 LTFEFSITIANQDR-----YLIGVDTQD--SLVYT---KTKNKFSTILKNLGYSVF 317
 DB 226 GAGVFNEAKAINERKSEADKRWVIGVDRDQDEGKTSKDGKEANFVLASSIKEVGKAVQ 285
 QY 318 SVLSDLTYTKSNSRNLAGFEFGKSAATVYIGIDRFVDIADTSLGNDKLTATEAISEAK 377
 DB 286 LINKQVADK-----FPGGK--TTVY-GLKDGVEIATY---NVSKEAVKAIKEAK 330

RESULT 5 US-10-451-337-40

Sequence 40, Application US/10451337
 Publication No. US20040097706A1
 GENERAL INFORMATION:
 APPLICANT: SHIRE BIOCHEM INC.
 APPLICANT: MARTIN, Denis
 APPLICANT: BRODEUR, Bernard R.
 APPLICANT: HAMEL, Josee
 APPLICANT: RIOUX, Stephane
 APPLICANT: RHEAULT, Patrick
 TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
 FILE REFERENCE: 12806-24PCT
 CURRENT APPLICATION NUMBER: US/10/451.337
 CURRENT FILING DATE: 2003-11-18
 PRIOR APPLICATION NUMBER: US 60/256,940
 NUMBER OF SEQ ID NOS: 42
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 40
 LENGTH: 330
 TYPE: PRF
 ORGANISM: S. pyogenes
 US-10-451-337-40

Query Match 10.0%; Score 216.5; DB 15; Length 330;
 Best Local Similarity 28.3%; Pred. No. 4.7e-08;
 Matches 102; Conservative 46; Mismatches 143; Indels 69; Gaps 17;

QY 39 AAKADANKHFGMLNMAIYTAGTVNDNSFNOSGMEAIQOLGALTG-GEITSVDSSTAIEG 97
 DB 19 ASKGAAGSGKTDLKAVAMVTDGVDKSFNOSAMEGLSQWKEMLQKGTGDFYQSTSES 78
 QY 98 KYSLANTNKNVWVLSGFQ--HGDAFTRWLKIPENKQILFTEKNIIILGIDMTDTEVNIPT 155

DB 79 EYA-----TNDTAVSGGYQLYIGIGA--LKDAIAKAGNBEKVFYIID----DIEGK 128
 QY 156 GRVINTLYKTEBAGMLAGVANNASFLAKKPSDPTKRSAL-VIGGISPATVDTPIAGYIAG 214
 DB 129 DNVAATVPADHEAAYLAGIAAIAK-----TTKTTGVPFGMEGTVITREKGEAG 179
 QY 215 IKAWNLKNSDKKTKITTDKIEINLGF----DVODSTKERLEQIASKDXPSTLLAVAGP 269
 DB 180 VKS-----VDDTQVAVDAGSFGDAKAGKTIAAQAAGAD--VIYQAAGGT 225
 QY 270 LTFEFSITIANQDR-----YLIGVDTQD--SLVYT---KTKNKFSTILKNLGYSVF 317
 DB 226 GAGVFNEAKAINERKSEADKRWVIGVDRDQDEGKTSKDGKEANFVLASSIKEVGKAVQ 285
 QY 318 SVLSDLTYTKSNSRNLAGFEFGKSAATVYIGIDRFVDIADTSLGNDKLTATEAISEAK 377
 DB 286 LINKQVADK-----FPGGK--TTVY-GLKDGVEIATY---NVSKEAVKAIKEAK 330

RESULT 6 US-10-451-337-41

Sequence 41, Application US/10451337
 Publication No. US20040097706A1
 GENERAL INFORMATION:
 APPLICANT: SHIRE BIOCHEM INC.
 APPLICANT: MARTIN, Denis
 APPLICANT: BRODEUR, Bernard R.
 APPLICANT: HAMEL, Josee
 APPLICANT: RIOUX, Stephane
 APPLICANT: RHEAULT, Patrick
 TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
 FILE REFERENCE: 12806-24PCT
 CURRENT APPLICATION NUMBER: US/10/451.337
 CURRENT FILING DATE: 2003-11-18
 PRIOR APPLICATION NUMBER: US 60/256,940
 NUMBER OF SEQ ID NOS: 42
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 41
 LENGTH: 330
 TYPE: PRF
 ORGANISM: S. pyogenes
 US-10-451-337-41

Query Match 10.0%; Score 216.5; DB 15; Length 330;
 Best Local Similarity 28.3%; Pred. No. 4.7e-08;
 Matches 102; Conservative 46; Mismatches 143; Indels 69; Gaps 17;

QY 39 AAKADANKHFGMLNMAIYTAGTVNDNSFNOSGMEAIQOLGALTG-GEITSVDSSTAIEG 97
 DB 19 ASKGAAGSGKTDLKAVAMVTDGVDKSFNOSAMEGLSQWKEMLQKGTGDFYQSTSES 78
 QY 98 KYSLANTNKNVWVLSGFQ--HGDAFTRWLKIPENKQILFTEKNIIILGIDMTDTEVNIPT 155
 DB 79 EYA-----TNDTAVSGGYQLYIGIGA--LKDAIAKAGNBEKVFYIID----DIEGK 128
 QY 156 GRVINTLYKTEBAGMLAGVANNASFLAKKPSDPTKRSAL-VIGGISPATVDTPIAGYIAG 214
 DB 129 DNVAATVPADHEAAYLAGIAAIAK-----TTKTTGVPFGMEGTVITREKGEAG 179
 QY 215 IKAWNLKNSDKKTKITTDKIEINLGF----DVODSTKERLEQIASKDXPSTLLAVAGP 269
 DB 180 VKS-----VDDTQVAVDAGSFGDAKAGKTIAAQAAGAD--VIYQAAGGT 225
 QY 270 LTFEFSITIANQDR-----YLIGVDTQD--SLVYT---KTKNKFSTILKNLGYSVF 317
 DB 226 GAGVFNEAKAINERKSEADKRWVIGVDRDQDEGKTSKDGKEANFVLASSIKEVGKAVQ 285
 QY 318 SVLSDLTYTKSNSRNLAGFEFGKSAATVYIGIDRFVDIADTSLGNDKLTATEAISEAK 377
 DB 286 LINKQVADK-----FPGGK--TTVY-GLKDGVEIATY---NVSKEAVKAIKEAK 330

Db 82 ESEYA-----TNLDPAVSGYQLIYIGIFA--LKDAIAKAGDNGVKFVIIIDDIIEGKD-- 133
 Qy 148 DTEVNIPTGVINULTYTEBAGMLAGYANASFLAKKPPSPPTKSAIVIGGISPATDF 207
 Db 134 ---NV-----ASVTFADHEAAVLAGIAAK-----TTTKRTVGFVGGEGVYTR 175
 Qy 208 IAGYLAGIKAMNLKNSDKTKITTDKIEINLGF-----DVQDSTKRLLEQIASDKPRST 262
 Db 176 FEFEGAGVKS-----VDDTIQVAVDVAGSGDAAKKTIAAAYAAAGAD--VI 221
 Qy 263 LLAVALPELTFESDIANON---DR-YLIGVDTDQ--SLVYT---KTKNKEFTSLKN 311
 Db 222 YQAAAGTGAGVFNBAINEKRESEADKVMVIGVDRDQDEKGYTSKDGKEANFVLASSIKE 281
 Qy 312 LGYSVFSVLDLYTKKNSRNLAGFEFGKKSATVYLGIDKRFVDIADTSLSEGNKXLA 371
 Db 282 VGRAVQIINKOVADK-----FPGGK--TIVY-GIKDGGVBIATT---NVSKEAVK 325
 Qy 372 AISEAKKEFEKTKTIP 388
 Db 326 AIKEAKAKIKSGDIKVP 342

RESULT 10

US-09-765-272-8
 ; Sequence 8, Application US/09765272
 ; Patent No. US20020061545A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccine
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/765,272
 ; FILING DATE: 22-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/961,083
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 328 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-09-765-272-8

Query Match 9.4%; Score 202; DB 9; Length 328;
 Best Local Similarity 25.7%; Pred. No. 5.9e-07;
 Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;
 Qy 33 RKSEIMAKADANHFGLNMAIVTAGTVNDSFNQSGWEAIIQOLGA---LTGSEITSVD 89

Db 3 RSSRNMASSSDVK-----TKAAIVTDTCGVDDKSPNQSAMGELQAWGKEHNLSDKNGFTYF 58
 Qy 90 SSTAELE-----GKYSLLANTNKNVWVLSGFQHGDAFTRWLKIPEKQOLFTEKNI 140
 Db 59 QSTSEADYANNLOQAAGSYNLIFGV-----GFALNNA-----VKDAKKEHTDLN 104
 Qy 141 ILGIDWDTENVIPTEGRYINULTYTEBAGMLAGYANASFLAKKPPSPPTKSAIVIGG 200
 Db 105 LIDVIRKQGNV-----ASVTFADNESGYLAGVAAAK-----TTTKQVGFVGGI 149
 Qy 201 -SPAVIDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVDQDSTKRLLEQIASDK 259
 Db 150 ESEVISRFEAGFRAGV-----ASVDSPIKQVYVAG-SFG-DAKAKKTIAAAYAAAGAD- 201
 Qy 260 PSTLLAVAGPL-TEIFSDIIANONDR-----YLIGVDTQSL--VYT-----KTKNKEFT 306
 Db 202 -IYQVAGGTGAGVFPEAKSLNESRPEKWKVIGVDRDQDEKGYTSKDGKESNVLV 259
 Qy 307 SILKNLGYSVFSVLDLYTKKNSRNLAGFEFGKKSATVYLGIDKRFVDIADTSLSEGNK 366
 Db 260 STLKQVGTIVKDI-----SNKAERGEFPGGQ---VIYYSJKDGVDLAVTNLSSEBK 308
 Qy 367 KLATEAISEAKKEFEKTKTIP 388
 Db 309 K---AVEDAKAKILDSGVKVP 326

RESULT 11

US-09-769-787-132
 ; Sequence 132, Application US/09769787
 ; Publication No. US20030091577A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Microbial Technics Limited
 ; APPLICANT: Gilbert, Christophe FG
 ; APPLICANT: Hansbro, Philip M
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: PWC/22129WC
 ; CURRENT APPLICATION NUMBER: US/09/769,787
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: GB 9816337.1
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: US 60/125164
 ; PRIOR FILING DATE: 1999-03-19
 ; NUMBER OF SEQ ID NOS: 388
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 132
 ; LENGTH: 350
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-09-769-787-132

Query Match 9.4%; Score 202; DB 10; Length 350;
 Best Local Similarity 25.7%; Pred. No. 6.4e-07;
 Matches 98; Conservative 55; Mismatches 145; Indels 84; Gaps 19;
 Qy 33 RKSEIMAKADANHFGLNMAIVTAGTVNDSFNQSGWEAIIQOLGA---LTGSEITSVD 89
 Db 25 RSSRNMASSSDVK-----TKAAIVTDTCGVDDKSPNQSAMGELQAWGKEHNLSDKNGFTYF 80
 Qy 90 SSTAELE-----GKYSLLANTNKNVWVLSGFQHGDAFTRWLKIPEKQOLFTEKNI 140
 Db 81 QSTSEADYANNLOQAAGSYNLIFGV-----GFALNNA-----VKDAKKEHTDLN 126
 Qy 141 ILGIDWDTENVIPTEGRYINULTYTEBAGMLAGYANASFLAKKPPSPPTKSAIVIGG 200
 Db 127 LIDVIRKQGNV-----ASVTFADNESGYLAGVAAAK-----TTTKQVGFVGGI 171
 Qy 201 -SPAVIDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLGFVDQDSTKRLLEQIASDK 259
 Db 172 ESEVISRFEAGFRAGV-----ASVDSPIKQVYVAG-SFG-DAKAKKTIAAAYAAAGAD- 223
 Qy 260 PSTLLAVAGPL-TEIFSDIIANONDR-----YLIGVDTQSL--VYT-----KTKNKEFT 306

Db 224 --IYYVAGGTGAGVFAEAKSLNESRPNENKRWVIGVDRDQAEAGKTYTSKDKESNFTLV 281
 QY 307 SILKNLGYSPVSLDIYTKKSNRNLAGFEFGKKSATVIGIDRFDIADTSLGNDK 366
 Db 282 STLQVGTIVDI-----SNKAERGFPGQ---VIVYSLDKGVDLVTLNISEBK 330
 QY 367 KLATEAISEAKKEFEKTKTIP 388
 Db 331 K---AVEDAKAKILLDGSVKVP 348

RESULT 12

US-09-071-035-80
 ; Sequence 80, Application US/09071035
 ; Publication No. US20020045737A1

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071.035

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB369P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-071-035-80

Query Match 8.3%, Score 178; DB 9; Length 339;

Best Local Similarity 25.1%; Pred.No.4e-05;

Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

Db 33 RKEIIMAKADANKHFGILNMAIVTAGGVNDNSFNQSGWEAIIQOLGA---LTGE----- 84

Db 7 KTAESGGGKGD-----AHSAVIITDTGVDKSFNQSWEGLQAWGKEHDLPEBSKGAY 62

QY 85 ITSVDSS--TALEGGKSSSLANTNKNVWVLSGFQHGDAFTRWLKIPEKNQLFTEKNITIL 142

Db 63 IQSDADADYTTNIDQAVSKFNITFGI-----GYLLKDA-----ISSAADANPDITNFVLI 112

QY 143 GIDMTDENVIPTGRYINLTYKTEBAGWLAGYANASFLAKKPPSDPTKRSATVIGGGISP 202

Db 113 DDQIDKGNV-----VSATFRDNDAAYLAGVAAA-----NETKTNKVGFGGEG 157

QY 203 AVTP-FLAGIAGIKAMLNKSDKTKTTTKIKIINLGFVDVDTSTKERLEQILSKDKPS 261

Db 158 VVDRFQAGEKGV-ADAAVELGKEITVFI-----KTAASFPADPA 196
 QY 262 TLIAVAGPLTEIFSDII-----ANQNR-YLIGVDTQSL--VYT 298
 Db 197 KQALAAAMVQNGVDIIIFHSGATGGVFPQEAQKLNESGDKVWVIGVDRDQADADKY- 255
 QY 299 KTK-----NKFPISILKNLGYSPVSLDIYTKKSNRNLAGFEFGKKSATVYIGIDRF 353
 Db 256 KTKGKEKDNFTLITSLTGVGTAVDIANRLEDK-----FPGSEHLVY-GLDDGG 304
 QY 354 VDIADTSLGNDKKLATEAISEAKKEFEKTKTIPAEVVKLTLEIPMPD 403
 Db 305 VDLTDGYL--NDKT-----KEAVTKAKDKVIGSDVAVPEKPE 339

RESULT 13

US-10-206-576-80
 ; Sequence 80, Application US/10206576
 ; Publication No. US20030017495A1

GENERAL INFORMATION:

APPLICANT: Choi et al.

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 497

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/206,576

FILING DATE: 29-Jul-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/071,035

FILING DATE: 1998-05-04

APPLICATION NUMBER: US 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: US 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: US 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB369P1D1

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 80:

US-10-206-576-80

Query Match 8.3%, Score 178; DB 14; Length 339;

Best Local Similarity 25.1%; Pred.No.4e-05;

Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

Db 33 RKEIIMAKADANKHFGILNMAIVTAGGVNDNSFNQSGWEAIIQOLGA---LTGE----- 84

Db 7 KTAESGGGKGD-----AHSAVIITDTGVDKSFNQSWEGLQAWGKEHDLPEBSKGAY 62

QY 85 ITSVDSS--TALEGGKSSSLANTNKNVWVLSGFQHGDAFTRWLKIPEKNQLFTEKNITIL 142

Db 63 IQSDADADYTTNIDQAVSKFNITFGI-----GYLLKDA-----ISSAADANPDITNFVLI 112

QY 143 GIDWTDRENYIPGRVYNLTYYKTEEAMTLGYNVNASPLAKKFPSPDXTKSALYIGGGISF 202

Db 113 DDQIDGKKNV-----VSATFRDNEAYIAGVAAA-----NETKTNKQFVGGEEG 157

QY 203 AVTD-PIAGYLAGIKAWNLKNSDKKITTTDKIEINIGFDVODTSTKERLEQIASHOKDS 261

Db 158 VVIDRFAGEGFEKV-ADAAKEIAGEITVDI-----KYAASFADPA 196

QY 262 TLLAVAGLPLEIFSDII-----ANQNR-YLIGVDTQSL--VYT 298

Db 197 KGRLLAAMQNGVDIIFFHSGATGCGVFEAKDNLBSSGGDKVWYIGVDRDDADGKY- 255

QY 299 KTK-----NKFSTILKNLIGSVFSLSDLYTKSKNSRMLAGEFPGKKSATYVLGIKDRF 353

Db 256 KTKGKGEDNFTLITSLGKGVTAQODINRALEDK-----PGGSHLYV-GIKDGG 304

QY 354 VDIADTSLBEGDDKMLATEALSEAKKEBEETKTIIPAEVRYKLTIEPMEWD 403

Db 305 VDLIDGVL--NDKT-----KEAVTKAKODKVISGDVKVPEKEE 339

```

1      RESULT 14
2      US-09-071-035-78
3      Sequence 78, Application US/09071035
4      Publication No. US20020045737A1
5      GENERAL INFORMATION:
6      APPLICANT: GIL H. Choi
7      TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
8      NUMBER OF SEQUENCES: 496
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: Human Genome Sciences, Inc.
11     STREET: 9410 Key West Avenue
12     CITY: Rockville
13     STATE: Maryland
14     COUNTRY: USA
15     ZIP: 20850
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
18     COMPUTER: HP Vectra 486/33
19     OPERATING SYSTEM: MSDOS version 6.2
20     SOFTWARE: ASCII text
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/09/071,035
23     FILING DATE:
24     CLASSIFICATION:
25     PRIOR APPLICATION DATA:
26     APPLICATION NUMBER:
27     FILING DATE:
28     ATTORNEY/AGENT INFORMATION:
29     NAME: A. Anders Brookes
30     REGISTRATION NUMBER: 36,373
31     REFERENCE/DOCKET NUMBER: PB369P2
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: (301) 309-8504
34     TELEFAX: (301) 309-8512
35     INFORMATION FOR SEQ ID NO: 78:
36     SEQUENCE CHARACTERISTICS:
37     LENGTH: 361 amino acids
38     TYPE: amino acid
39     STRANDEDNESS: single
40     TOPOLOGY: linear
41     MOLECULE TYPE: protein
42     US-09-071-035-78

```

Query Match 8.3%; Score 178; DB 9; Length 361;
 Best Local Similarity 25.1%; Pred. No. 4,4e-05;
 Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

Qy 33 RKSEIMAKADANKFGLINMAIVTGGTVNDNSFNQSCWEAIIQQTGA---LTGG----- 84
 29 KTAASGGGKGA---AHSAVIITDTGGVDKDSFNQSSWEGIQAGKHEHIDPEGSGYAY 84

```

OY      85  TTSVDS--TAELEKSYSLANTNGKMWVLSFGQGDAPFTRLKJPEKQKLTENKIIL 142
Db      85  IQSDNAAYTNIDDAVSKRTFFIGI---GYLKDA-----ISSAADNPDTNFI 134
OY      143  GIDWTDENVIPTRGYINLTYYKTEBAGMLAGYANASFLAKHPPSDPTKSAIVGGIS 202
Db      135  DDQIDGKGV-----VSATPRDNMAALVAGVAAA-----NETKNNYGVGGEGB 179
OY      203  AVTD-PIAGVLAGIKAMWLKNSDKTKITTDKIEINLGFVDVDTSTKERLEQIASDKPS 261
Db      180  VVIDRFQGFGEKV-ADAKELGKEITVDT-----KYAASPADA 218
OY      262  TLLAVAGLTFESPII-----ANQDR-YLIGVDTQSL--VYT 298
Db      219  KKKLAAAMYQNGVDIIFPHASGATGQGVFOEAKDLNBSGGKQWVIGVDRDODADGKY- 277
OY      299  KTK-----NKFFSTLKNLGVSVFVSVDLVTYKSNNSNLNLAGFEFGKKSATYVLGKDF 353
Db      278  KTKQGEKEDNFTLSTLKGVTAVODIARPALEDK-----PPGSEHLVY-GLKDG 326
OY      354  VDIADTSLBGDKKLATBAISEAKKEFEKTKTIPAEYVRKTLLEIPEMD 403
Db      327  VDLTDGVL--NDKT-----KEAVTADKQVIGSVKQVPEKRE 361

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match . 8.3%; Score 178; DB 14; Length 361;

Best Local Similarity 25.1%, Pred. No. 4.4e-05;
Matches 103; Conservative 51; Mismatches 140; Indels 116; Gaps 20;

```

Qy      33 RKSEIMAKADANGHFGINMAIVTAGTVNNDNSFNQSGWEAIIQOLGA---LTGGE----- 84
      29 KTAESGGGKGDA-----AASAVIITDTGVDKSFNQSWEGLQAMGKEHDLPEGSKGYAY 84
Db
Qy      85 ITSVDSS--TAELEGKYSSLANTNNKNNVVLGFGHGDAFTRMLKI PENKQLFTEKNIIIL 142
      85 IQSNDADYTTNIDQAVSSKFNITIFGI---GYLLKDA-----ISSNADANPDNFEVLI 134
Qy      143 GIDWTDTEENVIFPTGRYINLTFTKEBAGWLAGYANASFLAKKFPSPDPTKSAIVIGGGISP 202
      135 DDQIDGKKNV-----VSATFRDNEAAYLAGVAAA-----NETKTNKRVGVGGBEG 179
Db
Qy      203 AVTD-FLAGYLGIKAWMLKNSDKTKITTDKIEINILGFDVQDTSTKERLEQIASDKDPS 261
      180 VVIDRFQAGPEKGV-ADAKELGKEITVDT-----KYAASFADPA 218
Db
Qy      262 TLLAVAGPLEIFSDII-----ANQNDR-YLIGVDTQSL--VYT 298
      219 KGKALAAAMYONGVDIIFHASGATGGVFCQAKDLNESGSGDKYVIGVDRDQDADGKY- 277
Db
Qy      299 KTK-----NKEFTSILKNLGYSVFSVLSDLTYTKKSNRNLAGPFGKKSATVYLGIKDRF 353
      278 KTKDGKEDNFTLTSTLKGVGTAQODIANRALEDK-----FPGGEHLVY-GLKDDG 326
Db
Qy      354 VDIADTSLGNDKKLATEAISAKKEFEFEKTKTI PAEEVRKTLFEMPMD 403
      327 VDLTDGYL--NDKT-----KEAVKTAKDQVIGSDVAVPEKPE 361
Db

```

Search completed: December 18, 2004, 01:44:03
Job time : 98.7643 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 01:15:13 ; Search time 25.651 Seconds
(without alignments)
1586.669 Million cell updates/sec

Title: US-09-676-249D-4
Perfect score: 2155
Sequence: 1 MWDKETTKEKSDNQNKQI.....KQPDKQOESLDKLTIDNNL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
1	552	25.6	461	2	B90555 ABC transporter xy
2	203.5	9.4	350	2	F86804 basic membrane pro
3	202	9.4	350	2	G95097 lipoprotein (impor
4	198	9.2	374	2	C97965 conserved hypochet
5	192	8.9	357	2	D96986 probable lipoprote
6	181	8.4	353	1	H71340 membrane lipoprote
7	180.5	8.4	350	1	C70009 ABC transporter (l
8	178.5	8.3	359	2	F72418 basic membrane pro
9	177	8.2	341	2	E70147 basic membrane lip
10	173	8.0	357	2	AH1610 CD4+ T cell-stimul
11	168.5	7.8	356	2	C82946 hypochetrical prote
12	168	7.8	357	2	AD1248 CD4+ T cell-stimul
13	165.5	7.7	360	2	H70147 basic membrane pro
14	162.5	7.5	525	2	C82914 conserved hypochet
15	161	7.5	339	2	F70147 basic membrane pro
16	152.5	7.1	524	2	D64204 hypochetrical membr
17	150	7.0	524	2	D82944 hypochetrical membr
18	146	6.8	353	2	G70147 basic membrane pro
19	140	6.5	379	2	H75318 membrane lipoprote
20	139.5	6.5	337	2	AH2591 membrane lipoprote
21	139.5	6.5	337	2	E97374 Deinococcus radiot
22	136	6.3	547	2	E29504 mercury(II) reduct
23	134	6.2	539	2	D82886 conserved hypochet
24	132	6.1	326	2	G95857 hypochetrical prote
25	131.5	6.1	1558	2	B71603 R5A-H3 antigen PF
26	128	5.9	657	2	S73428 probable lipoprote
27	126	5.8	349	2	F84246 hypochetrical prote
28	125.5	5.8	763	2	A82863 hypochetrical prote
29	125.5	5.8	1223	2	E88451 protein K10D2.1 (l

30	123.5	5.7	626	2	C25035 Colicin Ia - Bache
31	122.5	5.7	322	2	F84236 ABC transporter (l
32	122.5	5.7	2346	2	T13829 Tpr homolog - Fru
33	120	5.6	1546	2	G90603 hypoprotein (impor
34	119.5	5.5	384	2	I40867 hypochetrical prote
35	119.5	5.5	2285	2	T12796 probable transglyc
36	118.5	5.5	556	2	H82301 peptide ABC transp
37	118.5	5.5	1140	2	S73786 hypochetrical prote
38	118.5	5.5	1963	2	B98002 Iga-specific met
39	118	5.5	1031	2	C81302 probable type I st
40	117	5.4	798	2	S62791 probable lipoprote
41	117	5.4	1959	2	AG1085 hypochetrical prote
42	116.5	5.4	396	2	AD2719 conserved hypochet
43	116.5	5.4	396	2	H97500 hypochetrical prote
44	115.5	5.4	350	2	F70139 exported protein (
45	115	5.3	282	2	C97271 probable xylanase/

ALIGNMENTS

RESULT 1

B90555 ABC transporter xyllose-binding lipoprotein [imported] - Mycoplasma pulmonis (strain UAB C)
C/Species: Mycoplasma pulmonis
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #ext_change 09-Jul-2004
C/Accession: B90555
R/Chamblaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A/Reference number: A99512; PMID:11353084
A/Accession: B90555
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-461 <RUR>
A/Cross-references: UNIPROT:Q980U5; GB:AL445566; P1D:G14089760; P1DN:CAC13519.1; GSPDB:GT
A/Experimental source: Strain UAB C11P
A/Genetics:
A/Gene: MYPU_3460
A/Genetic code: SGC3

Query Match 25.6%; Score 552; DB 2; Length 461;
Best Local Similarity 31.2%; Pred. No. 1.5e-26;
Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;

QY	13	ADNQNKOIT--DYSKISGLVNER-----KSEITMAKADANKFGLNMAITVAGTV 61
DB	28	AQNPNKTNMSLDSSKTIIDLSQKEVETQKIYENKIKQASLEFQK-----VLTITADGNI 82
QY	62	NDNSFNQSGWEATQOLGALTGEITS-----VDSSTALEGKYSSLANTNK 107
DB	83	DDISFNQGVYESQKTLKDFDYKAYKSONKEAENQHKLDNYINAVDLEQNYKVALDRGY 142
QY	108	NWVVLGSGFGHGDFTMLKLPENKQLFTEKNIIILIGIDWTDTEN-VIPTGRYINLYTKTE 166
DB	143	TTIILITFGQGNIEFNLNDENNLRPEKKNKVIIGDVAIPNNNSKIPQSLSLFTEKE 202
QY	167	EAGMLGYNASFLAKKFPSPDPTKRSALIVGGGISPATVDFIAGYLAGIRAMN--LKNSD 224
DB	203	EAGNQKIVASADLVGTYIANNEAKRAISAFGGDFACVTDPLFNGFEGDIANNSEANAN 262
QY	225	KKTITTDKLEINLGPVDVDTSTKERLEQJASRKQESTLLAVAGPLTEIFSDII---ANQ 281
DB	263	KKTKIVSENLVLTGF-IPNAENKEVSVNVERGKSTISLPVAGPFGVVDVLRKQTSO 321
QY	282	NDRLIVGVTDVDSLVYTKTKNKEFTSLKVLGYSVSVSLDYLT-----KKSNSRNL 333
DB	322	EDFFIVGVDPDQSLFSTNDSKREFTSLVKNIAFPVQIILALLTKQESVILKGGDKFL 381
QY	334	AGFEFGKASATVVLGIDRFVDIADTSLKNDKKLATEAISEA--KKFEKTKTTPAE 391
DB	382	GS--NPKNLVLRKGLSAKFVNITTKSVKESITQADTSIQKALDKMANPNNSKLIKEM 438

QY	392 VRKTL	397
Db	439 TNGDL	444

RESULT 2

basic membrane protein A [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86804
R:Botolin, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86804
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <STD>
A:Cross-references: UNIPROT:O9CFM9; GB:AE005176; PID:gl2724428; PIDN:AAK05536.1; GSPDB:C
A:Experimental source: strain IL1403
C:Genetics:
C:Gene: bmpA

Query Match	9.4%	Score 203.5;	DB 2;	length 350;
Best Local Similarity	24.7%	Pred. No. 2.7e-05;		
Matches 97;	Conservative 63;	Mismatches 157;	Indels 75;	Gaps 17;

```

Qy 20 ITTVSKISJGVNERKEIMAAKADANKHFGMLMAITVTAAGTVYNNNSFNQSGWEALQDGA 79
Db 13 LASVAAYLAGCRSHDAAGSGRAKTD-----LMAALVTEIGGVNDRSFQSGMBEGLQSGWK 66
Qy 80 LTG-----GETTSVDSSTAELBGRKYSLSLANTKNWVLSGCFOHGGAFTRMWIKIPENQOLF 134
Db 67 ENNLKKGDTGYTVAQNSASADYTTNNNSAAGQGYKLLFPTGSGLSDDATISAALK---NN--- 120
Qy 135 TEKNIIILGIDWTDTEENVIPTRGRYINLTYKTEBACMLAGYANASFLAKKPPSDPTKRSAI 194
Db 121 PKSNFVYIVDSVIMDKQKV-----ASATPADNESAYLAGVAALK-----ATKTKNKI 165
Qy 195 -VIGGGISPAVNDPFIAGLYLAGIKANMLKNSDKKTKITTDKIBILNGPVDQPTSTKHELEQ 253
Db 166 GFIFGQOSDVITTRFEKGYEAGAKSVN-----PDIKVDVYAGSFSDAAKGITIAA 215
Qy 254 IASKDPRSTLLAVAGPL-TEIFSDIIA-----NQNDR-YLFGVDTDSLV--YT-----KT 300
Db 216 AMTGAAGDDVYQCAAGVGCTGVSEAKALNSTNEADKMWIVGVDDODEYLGKTYKSKQGD 275
Qy 301 KKNFETSLKNLGVSVFSVLSDLYTRKGSNSRNLAC-EEFGKSAAYYLGIKORFYVDIAD 358
Db 276 SNFVLVSTIKEVG---NVVKDIADKTDGKRGPGTITVYLDKNGGVNLGL-----DSAN 326
Qy 359 TSLGNDKRLATEALSEAKKEFEERTKTIYPAE 390
Db 327 SEIK-----DAVAKAKADIIDGKITVPSK 350

```

RESULT 3

1|proprotein [imported] - Streptococcus pneumoniae (strain TIGR4)
 C|Species: Streptococcus pneumoniae
 C|Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 A|Accession: G95097
 R|Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
 nson, T.J.; Hickney, E.R.; Holt, I.E.
 Science 293, 498-506, 2001
 A|Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A|Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A|Reference number: A95000; MUID:21357209; PMID:11463916
 A|Accession: G95097
 A|Status: preliminary
 A|Molecule type: DNA

A:Residues: 1-350 <KUR>
A:Cross-references: UNIPROT:Q97RH0; GB:AE005672; PIDN:AA74976.1; PID:q14972319; GSPDE:G14972319
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0845
C:Superfamily: ABC transporter ynfN

Query Match	9.4%	Score 202;	DB 2;	Length 350;
Best Local Similarity	25.7%;	Pred. No. 3.3e-05;		
Matches	98;	Conservative	55;	Mismatches 145;
			Indels	84;
			Gaps	19

```

QY      33 RKSEIMAAKADANKHFGMLMAIVTLAGTINDNSFNOSGEAIIQOLGA---LTGEIITSVD  89
Db      25 RSRRNAASSSDVK---TKAAIVTDGTGDVDSFNOSAMEGLQAMGEHNLSKONGFYTF  80
QY      90 SSTAELE-----GKYSILANTNKNWVLSGFOHDAFTFWLKIPENKQLEFTEKNII  140
Db      81 QSTSEADYANNLQQAAGSYNLIIFGV-----GFLNNNA-----YKDAKHEHTDLNIV  126
QY      141 ILGIDMTDTENVIPGGRYINLTYKTEEAGWLAGYANASFLAKKFPSPDPTKSAIVIGGI  200
Db      127 LIDVIYKQKQNV-----ASVTFADNESGYLAGVAAMAK-----TTKTKQGVFGGI  171
QY      201 -SPAVTDPLAGLAGIKAMNLKQSDKKTKITTDKIEINLGFVDVDTSTKERLEQIASDK  259
Db      172 ESEVISREAPGPKAGV-----ASVDPSIKVQVDYAG-SFG-DAKGMTIAAAQQAAGAD-  223
QY      260 PSTLLAVNGPL-TEIFSDIIANQNDR-----YLIGVDTQSL--VYT-----KTKNKPFT  306
Db      224 --IYQVANGGTAGYVFAEAKSLNESHPENEKWWVIGVDRQDEABGKTTSKDGKSNFVLV  281
QY      307 SILKNIAGYVSFVSLDLYTKKSNRNLAGFEFGKKSATVYLGIKDRFVIDIADTSLLEGNDK  366
Db      282 STLKQVGTIVKDI-----SNKARERGEPPGQ---VIVYSLDKQKVDLAVTNLSEBGK  330
QY      367 KLATEAISEAKKEPEEKTKTIP  388
Db      331 K-----AVEDAKAKIILDSGVKVP  348

```

RESULT 4

conserved hypotoxinical protein spr0747 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-oct-2001 #sequence_revision 22-oct-2001 #text_change 09-Jul-2004
C:Accession: C97965
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; Dehoff, B.S.; Edwards, R.; LeBlanc, D.O.; Lee, L.N.; Letkowitz, E.O.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
Y: Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balez, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; PMID:11544234

Query Match 9.2%; Score 198; DB 2; Length 374;
 Best Local Similarity 25.4%; Pred. No. 6.4e-05;
 Matches 97; Conservative 54; Mismatches 147; Indels 84; Gaps 18

Cy 33 RKSSIMAKKDANHPGLMNAITVAGGVNDNSNOSWEALQQLGA---LTGGSEITVD 89
 ::: | | | | |
Db 49 RSRNNMASSSDVK---TKAALTVDGTGVDKSDQSMBEGLQMGKEHNLSKDNGFFYF 104

Cy 90 SSTAELE-----GRYSLSLANTKNRWVLGSFGHDATFRMLKIPENKOLFTEKKII 140
 ::: | | | | |
Db 105 QSTEADYANNILOOASSYNI-----FVGALHNVAEVKKEHTDLNVY 150

141 ILGIDWTDNTNVIPTGRYINLTYTEBAGLAGYANASFLAKKPPSPPTKRSIVICGGI 200
151 LIDVDVADQKQV-----ASVTFADNBSGYLAGVAAAK-----TTRTKQVGFVGGI 195
201 -SPAVNDFIAGYLAGIKAMNLKNSDKTKITTDKIEINLIGFDVODTSTKERLEBIDAKDK 259
196 ESEVISEFEEGFKAGV-----ASVDPSTIKQVQVADAG-SFG-DAAKGKTIAAQAAD- 247
260 PSTLLAAGPL-TEIFSDIANQNDR-----YLIGVDTPQSL--VYT---KTGNKFEFT 306
248 --IVYQVAGGTGACVFAEASLSNBSRENKRWIGVDRDDEAGKTTSDGKESNVLV 305
307 SILKNLGYSVFVSLDLYTKKNSRNLAGFEGKKSATVYLIGDRFVDIADTSLGENDK 366
306 STLKQVGTVMKI-----SNVAKEGFPGG---VIYVSLDKQVDLAVTLVLSSEBK 354
367 KLAETAISEAKKEFEKTKITP 388
355 K---AVEDAKAKALIDGSAKVP 372

RESULT 5

probable lipoprotein, Med/BMP family [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: D96986
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: D96986
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <KUR>
A/Cross-references: UNIPROT:Q97L60; GB:AE001437; PIDN:AAK78679.1; PID:G15023581; GSPDB:C
A/Experimental source: Clostridium acetobutylicum ATCC924
C/Genetics:
A/Gene: CAC0702

Query Match 8.9%; Score 192; DB 2; Length 357;
Best Local Similarity 23.3%; Pred. No. 0.00014;
Matches 94; Conservative 67; Mismatches 157; Indels 86; Gaps 18;
20 ITDVSISGLVNEKSEIMAA--KADANKHGLMAIVTAGCTVNDSPNOSGWEALQOL 77
10 LTTVMIVAGLPAGCSSTSSGSGNSKDTK--VKVGLSTDEGGLNDSFNGADEGIIKA 66
78 GALTGEITVDSSTA-ELBGKYSLLANTNKNVVLGSGFHGDAF-TRMLKIPENKOLFT 135
67 AKESYVDVYKAIIESKKDDYQPNLOSLDNDSDLVFGVGYQMAADLAIATKYPKPKAII 126
136 E-----KNIILIGIDWTDNTNVIPTGRYINLTYTEAGWLAGYANASFLAKKPPSDP 188
127 DDVAVDQPKNI-----MSLVFKEQSGFLMG-----VIAKGM----- 158
189 TKRSAL-VIGGISPATDPIAGYLAGIKAMNLKNSDKTKITTDKIEINLIGFDVODTST 247
159 TKTNKIGFVGKQOPLNKPLKLSGIYAGAKTVN-----PNITVEK---NTNYSPTS 208
248 -KERLEBIASDKRSTLLAVAGPLTEIFSDIANQNDR-----YLIGVDTPQSLVYTTK 302
209 GKEVATSLVNGCCIVVHAAGAGIYV-FVAKSLRQGDVMAIGVDKQQAALPKYAD 267
303 KFTTSLIKNLGYSVFVSLDLYTKKNSRNLAGFEGKKSATVYLIGDRFVDIADTSL 362
268 VILTSWVKGVDIATYNTVLDLVKGGK-----FEGKVES---FGKEDGVAVPTS-- 315
363 GNDKLAETAISEAKK-----EFEKTKITIPAEVVR 393
316 -NKHVPSEVLSTLVDRYKKAIIIDGKIVPPTVDAQTFKTDQIK 357

RESULT 6

H71340
membrane lipoprotein tmpc precursor - syphilis spirochete
C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C/Date: 07-Aug-1998 #sequence_revision 07-Aug-1998 #text_change 09-Jul-2004
C/Accession: H71340; A43595; S29561
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uetexback, T.; McDer
they, L.; Weidman, J.; Smith, H.O.; Vener, J.C.
Science 281, 375-388, 1998
A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A/Reference number: A71250; MUID:98332770; PMID:965876
A/Accession: H71340
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-353 <COL>
A/Cross-references: UNIPROT:P29724; GB:AE001211; GB:AE000520; NID:G3322582; PIDN:AA6530;
A/Experimental source: strain Nichols
R;Schouls, L.M.; van der Heide, H.G.J.; van Embden, J.D.A.
Infect. Immun. 59, 3536-3546, 1991
A/Title: Characterization of the 35-kilodalton Treponema pallidum subsp. pallidum recombi
A/Reference number: A43595; MUID:91372962; PMID:1894360
A/Accession: A43595
A/Molecule type: DNA
A/Residues: 1-10, 'A', 12-158, 'R', 160-353 <SCH>
A/Cross-references: GB:X57836; NID:948838; PIDN:CAA40968.1; PID:G581809
A/Note: this protein is shown to incorporate palmitic acid
C/Genetics:
A/Gene: tmpc; TP0319
A/Start codon: GTG
C/Superfamily: ABC transporter ynfN
C/Keywords: blocked amino end; lipoprotein; membrane protein; thiolester bond
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-353/Product: membrane lipoprotein tmpc #status predicted <MAT>
F:21/Modified site: fatty acylated amino end (Cys) (in mature form) #status predicted
F:21/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status predicted

Query Match 8.4%; Score 181; DB 1; Length 353;
Best Local Similarity 23.7%; Pred. No. 0.00065;
Matches 86; Conservative 69; Mismatches 134; Indels 74; Gaps 17;

52 MAIVTAGTVNDSPNOSGWEALQOLGALTGEITVDSST-ALBGKYSLLANTNKNV 110
42 VGMVTDGSDIDDKSFNQVVEGISRFAQENNAKCKVTASTDAEYVPSLSAFADENMG 101
111 VLSGFQHGDAFTWMLKIPENKOLFTENKIIILIGIDWTDNTNVIPTGRYINLTYTEAGW 170
102 VAGGSFLVEA-----VIETSAFPPKQFLVIDAVVODRNV-----VSAVFGQNGSF 149
171 LAGYANASFLAKKPPSPPTKRSAL--VIGS--GISPAVDFIAGYLAGIKAMNLKNSDK 226
150 LVGYA-NALAKA-----AKSAVGFIVGMEGLMPL--FEGAFEGVAVD----- 193
227 TKITTDKIEINLIGFDVODT-STKERLEQIASK--DKPSTLLAVAG---PLTEIFSDII 278
194 -----PDIVVVEVANTSPDPKQGOALAAKLYDSCVNIPOVAGGTGNGVIEKARR 246
279 ANQNDKYLIGVDTPQSL--VYTTKAKKFTTSLIKNLGYSVFVSLDLYTKKNSRNLAGF 336
247 LNDQDVVIVIGVDDQVMDGYDYSKSVLTSWVRA-----DVAAEKISKAADGSPG- 300
337 EFGKKSATVYLIGDRFVDIADTSLGENDKLAETAISEAKKEFEKTKITIPAEVVKTL 396
301 -----GQSIMFGLEDKAVGLPE-----ENPNLSAVMEKIRSEFKI-----VSK 342
397 EIP 399
343 VVP 345

RESULT 7

C70009

R:Ojaimi, C.; Davidson, B.E.; Saint Giron, I.; Old, I.G.
Microbiology 140, 2931-2940, 1994
A:Title: Conservation of gene arrangement and an unusual organization of rRNA genes in
A:Reference number: 140241; PMID:95111614; PMID:7812434
A:Accession: 140242
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-44, 'A', 46-179 <RE2>
A:Cross-references: GB:135050; NID:9516591; PIDN:AAC41402.1; PID:9551744
C:Superfamily: basic membrane protein C

Query Match 8.2%; Score 177; DB 2; Length 341;
Best Local Similarity 26.7%; Pred. No. 0.0011;
Matches 100; Conservative 59; Mismatches 124; Indels 92; Gaps 25;

QY 44 ANKPGFLMAIYAGGVNDNSFNQSGWEALIQGLTGEELTSVNSTALEGKYS-- 101
DB 24 SSKKIKISMLV---DGLVDDKSFNSSANLRLKDPENIEVFS--CAISGVSSYV 78
QY 102 --LANTNKN---VWVLSGFQHGDAFTRLKIPENKOLFTEKNII--ILGIDWTDENVI 153
DB 79 SDDNLKRNSSDILW-LVGMILDA--SLVSENPKI--SYGIDPIYDDVQIPENLI 133
QY 154 PIGRYINLYKTEBAGMLAGY--ANASFLAKFPSPDFTKSAIVIG--GSIPTAVTD-FI 208
DB 134 A-----VVRVREGAFLAGYIAKKSFSGK-----IGFICKMKNIVDAFR 174
QY 209 AGYLAGIKAMNLKNSDKKITTIDKIEINLGFVQDSTKERLEQIASK--DKPSTLLAV 266
DB 175 YGESAKYKAN-KDIEIISYNSFSFDVIG-----RTLAKSYSGKIDVIF 221
QY 267 ACPLTEI-FSDLIANQND-RYLIGVDTQSLVYTKTKNKFETSLKNLGYSFVSLSDLV 324
DB 222 AAGLAGIVETAKNLGDDYVIGADDD--YLAPEKN-FITSVIKIKGALVITGE-Y 277
QY 325 TKKSNRNLAGFEFGKKSATVYIGIKDRFVDIADTSLGNDKKLATEAISAKKEFEKT 384
DB 278 IKNNVY-----WEGSK-----VWQGLRDGVIGLPMAN-----EF-EYI 310
QY 385 KTIPEAEVKTEIIP 399
DB 311 KYLERKIKNEIIVP 325

RESULT 10

CD4+ T cell-stimulating antigen, lipoprotein [imported] - *Listeria innocua* (strain Clp1)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1610
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fahl, H.; Jones, L.M.; Karsch, U.
Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schleuter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A.:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <GLA>
A:Cross-references: UNIPROT:Q92BW7; GB:AL592022; PIDN:CAC96656.1; PID:G16413898; GSPDB:C
A:Experimental source: strain Clp11262
C:Genetic:
A:Gene: tcsA
C:Superfamily: ABC transporter yuIN

Query Match 8.0%; Score 173; DB 2; Length 357;
Best Local Similarity 24.2%; Pred. No. 0.0021;
Matches 101; Conservative 61; Mismatches 138; Indels 118; Gaps 23;
QY 27 SGLV-----NKRSEIMAAKANKHFLNMAIVTAGTVNDNSFNQSGWEALIQGL 78

DB 16 SGLVLAGCGSSSDDKS-----GDKKSKDFTVAWTTDGVDDRSFNQSAWEGIQKFG 69
QY 79 ALV-----GGEITVDSSTALEGKYSLLANTNKNVWVLS-----GFOHGDAFTRLK 126
DB 70 KANDMEKTDGIVYLOSASEADYK-----TLNLTVRSYDYLIGYIKLAIIEVSK 123
QY 127 -IPEKNOLFTEKNIIILGIDWTDENVIPGRYINLYKTEBAGMLAGVANSFLAKFP 185
DB 124 QPKNQFAIVDDTI-----DDRNVVSG-----FKNDOSYLVGVAG----- 163
QY 186 SDPTKSAIVYGGGIPAVTD-FIAGYLAGIKAMNLKNSDKKITTIDKIEINLGPVQ- 243
DB 164 --TTKTKNKFVGVGVAVIDREAGFTGVKA-----VNPQAQIDVQY 205
QY 244 --DTSTKERLEQIASKDS-----TLAAVAGPLTEIFSDLIANQND-----YLIGVDT 291
DB 206 ANDPAKADKQQLASSMSYSSGVVITHAAGTNGVFAE-AKQLKKDPSRAVWVIGVR 284
QY 292 DQ-----SLVYTKTKNKFETSLKNLGYSFVSLSDLVTKKSNRNLAG-FEFGKKSAT 344
DB 265 DQWDEKGVANDGKYNVTLTSEIKKVDIAV-----DLAT-----RTKADFPGGTK--- 312
QY 345 YVIGIKDRFVDIADTSLGNDKKLATEAISAKKEFEKTKTIPEAEVKTEIPEIP 402
DB 313 IEYGLDKDAVGLSE--HQNISKDVLAKE--EYKQIKVDG-----IKVPEKP 357

RESULT 11

C82946
hypothetical protein U0012 [imported] - *Ureaplasma urealyticum*
C:Species: *Ureaplasma urealyticum*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: C82946
R:Glaser, J.I.; Leikowitz, E.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of *Ureaplasma urealyticum*: Alternate views of a min
A:Reference number: A82870
A:Accession: C82946
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <GLA>
A:Cross-references: GB:AE002100; GB:AF222894; NID:G6898946; PIDN:AAF30417.1; GSPDB:GN001;
A:Experimental source: serovar 3; biovar 1
C:Genetic:
A:Gene: U0012
A:Genetic code: SGC3

Query Match 7.8%; Score 168.5; DB 2; Length 516;
Best Local Similarity 25.1%; Pred. No. 0.0066;
Matches 110; Conservative 65; Mismatches 167; Indels 97; Gaps 26;

QY 48 FGLNMAIVTAGGVNDNSFNQSG--WEALIQGLALGGETTSVDSSTAE-LEKYSSTLAN 104
DB 17 FALGIIIVATSCSTOKSTLVYSGFYW-----TSPSDDEGFTKYSMAS 62
QY 105 TNNVWVLSGFQHGDAFTRLKIPENKQ--LFTK--NIIIV--GIDWTDENVIPGT 156
DB 63 DGRALIMPGYQI-----PERLQNALVNDKDPNLIALLDGVYNNDKAKERYK 112
QY 157 --RYINLYKTEBAGMLAGVANSFLAKK--PSPDFTKSAIVIGG--GISPA-VTFI 208
DB 113 ADVADVDFVDAAFGLGIAAAYMLNSQAVGAD--NKLTWGGYVGINAKTTNYL 168
QY 209 AGYLAGIKAMNLKNSDKTK--ITTDKIEIN-----GFPVQDSTKERLEQI 254
DB 169 AGFDLGVKANETLKDKNKIKQESTQETKWINVEQYVASESSAGGQSDSANAKKIIQEL 228
QY 255 ASKQDSTLLAVAGPLTEI-FSDLIANQNDRY-LIGVD-----TDQSLVYTKTKNKFETSI 308
DB 229 ITKG-ADLIIPVAIPQVGLVTAIATTSINNVIGVDVIEINDQAI--NKTKDPFINTH 285
QY 309 L--KNLGYSVFVSLS--DLVYTKKSNRNLAGFEFGKKSATVYIGI---KDRF---VDI 356

Db | | | | | : | | | | | : | | | | | : | | | | | :
286 LSGMKN-GVIRFSTTKRLDTATILLENALKESLSKSDQIVIGSEIDPKDKYKLGNT 344
Qy 357 ADTSEF-----NDKKLATEAISAEKKEFEKKTTPAEVRKLTLEIPEMP 403
Db 345 VGNISDGVGVGSPSAHYVIDAFMLAQTNSDKYSTDELVNKLTNDLFTLTKKPYVD 404
Qy 404 KQDPKQESLDKLTIDINN 422
| | | | | : | | | | | : | | | | | : | | | | | :
Db 405 GYLDVKKET-DNEASLTNN 422

RESULT 12
AD1248
CD4+ T cell-stimulating antigen, lipoprotein [imported] - *Listeria monocytogenes* (strait
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1248
R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Me
ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1248
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <GLA>
A/Cross-references: UNIPROT:Q46754; GB:NC_003210; PIDN:CAC9466.1; PID:G16410817; GSPDB:
A/Experimental source: strain BGD-e
C/Genetics:
A/Gene: tcsA
C/Superfamily: ABC transporter yuFN

Query Match 7.8%; Score 168; DB 2; Length 357;
Best Local Similarity 23.8%; Pred. No. 0.0042;
Matches 98; Conservative 62; Mismatches 145; Indels 106; Gaps 22;
Qy 26 ISGLVNRKSEIMAKADANKHFGMLNMAIVAGTVVNDNSFNQSGWEALQGLALT----- 81
Db 19 ILGACSSSDPKKSSDDKSKDP--TVAMVDTGQVDRSFQSGAMEGLQFGKANDMEK 76
Qy 82 GGEITVSSTAELEGGYSSSLANTNKNVWVLS-----GFQHGDAFTRWK-IPEKKQ 132
Db 77 GTDGVNVLQSGSEADYK-----TNLNTAVASVDVLTLYGKXKDAIEVSKQKPKNQF 130
Qy 133 LFTFKNIIILIGDWTDBNVIPTRVYNLYKTEBAGMLAGYANASFLAKKFPSPDPTKRS 192
Db 131 AIVDDT-----DHRDNVVSIG-----FKDNDGSLVGVVAGL-----TTKTN 168
Qy 193 AIVGGGSPAVTD-FTAGYLAGIKANMLKNSDKKTKITTDKILINIGFDVQ---DSTK 248
Db 169 KVGCVGKGTVIDRFGHFTAGVKA-----VAPNQIDQVANDPAKA 212
Qy 249 ERLEQIASKQPS--TLAVAGPLTEIFSDIIANONDR-----YLIGVTDQ----- 293
Db 213 DKGGQIASMWSGVVDVIFPAAGGTGNGVFAE-AKMLKKQPSRAVWITGVRQWMBEGK 271
Qy 294 -SLVYTTTKKFFSIILKNLGYSVFVSLDLYTKKSNRLAG-FEFGKKSATYVLGIDK 351
Db 272 VTANDGQDVNLTLSIKRVDIAV---EDLAT-----RAKAGDPFGGT---LEYGDK 319
Qy 352 RFVDIADTSLGNKKLATEAISAEKKEFEKKTTPAEVRKLTLEIPEMP 402
Db 320 DAVGLSE-----HODNISKDVLAKE--EYKQKIIVGD-----IKVPEKP 357

RESULT 13
H70147
basic membrane protein D (bnpd) homolog - *Lyme disease spirochete*
C/Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 29-Sep-1999
C/Accession: H70147
R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
son, D.; Peterson, J.; Keriavage, A.R.; Quackebush, J.; Salberg, S.; Hanson, M.; Vugt,
.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
A/Reference number: A70100; MUID:98065943; PMID:9403685
A/Accession: H70147
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-360 <KLE>
A/Cross-references: GB:AE001144; GB:AE000783; NID:92688291; PIDN:AAB91505.1; PID:9268829
A/Experimental source: strain B31
C/Superfamily: basic membrane protein C

Query Match 7.7%; Score 165.5; DB 2; Length 360;
Best Local Similarity 25.1%; Pred. No. 0.006;
Matches 87; Conservative 64; Mismatches 138; Indels 57; Gaps 20;
Qy 37 IMAKADANKHFGMLNMAIVAGTVVNDNSFNQSGWEALQGLALTGGEITVSSTAELE 96
Db 34 VACSSSDGSEAKTVSLI-VDGAFDQKFNSSSKAIRLKDNLNITIEKSTGNSYL 92
Qy 97 GKXSLANTNKN-VWVLSGFQHGDAFTRWK-IPEKKQILIGDWTDBNVIP 155
Db 93 GDIANLEDGNSNLIWGI-GRSLDIL--FQRASENVSV---NYAI--IGVVDIPIK 143
Qy 156 GRYINLYKTEBAGMLAGYANASFLAKKFPSPDPTKRAIYIGGISAIVND-PIAGYLAG 214
Db 144 -NLINISFRSEEVAFIAGY-----FASKASKTGKIGFVGGRKVLSEFMYGEAG 193
Qy 215 IKANLEKNSDKTKITTDKI-----EINIGFDVDTSTKERLEQIASKDKESTLLAVAGPL 270
Db 194 AKYAN-----SNKVVSGYVGTGDFGLG-----RSTASNM---YDGVDIIFAAG-L 238
Qy 271 TEIFSDIIANQ--NDRYLIGVDTDQSLVYTKNKFSTILKNLGYSVFVSLDLYTKS 328
Db 239 SGIGVIEAAKELSPDHYIIGVDQSYL--APNNVIVSAVKKVDLSIMVSL-----TKKY 290
Qy 329 NSRNLAGFERGKKSATYVLGK--DRFVDIADTSLGDKKLATLEAI 373
Db 291 LETGV--LDGSK--TWFLGLKEDGLGLVLENMLKNSYSEIYNKSL 331

RESULT 14
C82914
conserved hypothetical membrane lipoprotein U0226 [imported] - *Ureaplasma urealyticum*
C/Species: *Ureaplasma urealyticum*
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: C82914
R/Glaser, J.I.; Leftkowitz, B.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A/Description: The complete sequence of *Ureaplasma urealyticum*: Alternate views of a min
A/Reference number: A82870
A/Accession: C82914
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-525 <GLA>
A/Cross-references: GB:AE002123; GB:AF222894; NID:96899229; PIDN:AAF30635.1; GSPDB:GN001;
A/Experimental source: serovar 3; biovar 1
C/Genetics:
A/Gene: U0226
A/Genetic code: SGC3

Query Match 7.5%; Score 162.5; DB 2; Length 525;
Best Local Similarity 22.9%; Pred. No. 0.016;
Matches 95; Conservative 63; Mismatches 138; Indels 119; Gaps 20;
Qy 99 YSGSLANTNKNVWVLSGFQHGDAFTRWK-IPEKKQILIGDWTDBNVIP 156
Db 72 YASKIQDGAHMLGLISFRHKNPISKYFNSPKDQ--QVSAVLI-----DEIYDLQIGKD 123

```

QY 157 RYILUYKTEAGLAAVANAASFLAKK---PSPSPTKRSALVIGGGS---PATYDPIAG 210
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 RIASITIRADQAAFLAGIAAAYLYNSNONVFGKD---NKLTWGGFPGIHLPSITRIIOG 179
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 YLAGIKAMNLLKNSDKKTKITTDKIE--INL-----GFVDQDSTYKER--LEQIAS 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 FKFEIQANELKKNKYKQKTEENNEKEMINEQVAFATNYQSGDPSPIPSDAXAIWNLVS 239
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 KDKPSTLLAAGPLTEIFSDIIANQNDR-YLIGVDTQSLVYTKNKKFP--TSILKNLG 313
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 -NNVDLILPVAQPIDYATTAASRSPIVVGVDTEQELDDNTNKARISENNKSLANGK 298
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 YSVSVSVS--DLYTK-----KSNRNLAGFERGKKSATVYLGIDRPFYDIADTS 360
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 TIISIVYRDLDFALFKGALLKASBGAQITNDINKDAYKLGHTTEASF--NKNITYVD--NTA 354
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 L-----EGNDKKLATEAISE-----AKKEFEKTKTIP 388
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 LVLELSKAGHOYLIDAIKLSGLKEVNDYKTVIEIIOBDFLLLSQIGTKLDEVAITSQO 414
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 389 AEEVRKTLLEIPEMP-----DKOPDKQOSES 412
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 415 GDWVLKS-EYGDLPFIQLOKMLGGLVYVDQKNELPYELSNSFLYLEKDPNKRQAS 468
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

basic membrane protein A (bmpa) - Lyme disease spirochaete
N:Alternate names: antigen P39; membrane lipoprotein A
C:Species: *Borrelia burgdorferi* (Lyme disease spirochaete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: F70147; I40289; I40241
C:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugr
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70147
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-339 <RES>
A:Cross-references: UNIPROT:Q45010; GB:AE001143; GB:AE000783; NID:G3688279; PIDN:AA667575
A:Experimental source: strain B31
R:Simpson, W.J.; Cieplak, W.
FEMS Microbiol. Lett. 119, 381-388, 1994
A:Title: Nucleotide sequence and analysis of the gene in *Borrelia burgdorferi* encoding t
A:Reference number: I40289; MUID:94327086; PMID:8050720
A:Accession: I40289
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-339 <RES>
A:Cross-references: GB:U24194; NID:G508420; PIDN:AAA72406.1; PID:G508421
R:Ojalmt, G.; Davidson, B.E.; Saint Glrons, I.; Old, I.G.
Microbiology 140, 2931-2940, 1994
A:Title: Conservation of gene arrangement and an unusual organization of rRNA genes in t
A:Reference number: I40241; MUID:95111614; PMID:7812434
A:Accession: I40241
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 248-339 <RE2>
A:Cross-references: GB:U5050; NID:G516591; PIDN:AA41401.1; PID:G516592
C:Genetics:
A:Gene: bmpa
i:Superfamily: basic membrane protein C

```

Db      37  GTFDDKSNESALNCVKYKVEBEFKIELVLKESSSNSYSLDLBG-----LNDAGSDDLTLWLG 92
Oy      115  FOHGD-AFTRMWLPKIPENKOLFTEKNIILIGIDMTDTEVNIPTGRYINLTYKTEEAGWLAG 173
Db      93  YRFSIVAKVALQNDMKYAL-----IDPIYSNDIP-ANLVGMTFRAQEGAFLTG 142
Oy      174  YANNSFLAKKEPPSDPKRSALVIGGIGISPAVTD-FLAGYLAGIKAMWLKNSDKKTKITTD 232
Db      143  YIAAKL-----SKTKIGIGFLGIBGEIYDAFRGYEAAKAN-----KDINISTQ 188
Oy      233  KI-----EINLGPYODSTKERLEQIASKDKSTLLAVGPLETPEISDIIRANO--NDRYL 286
Db      189  YIGSFADIEAGSVATRMYSDEID-----IHHAGIGGIGAIIVAKELSGCHYI 238
Oy      287  IGVDPDGLVYTKTKNQKFEFTSILKMLGYSVFSLDLYTKKNSRMLAFEFEGKSATVY 346
Db      239  IGVDEDAVYL---APDNVITSTTKVGRAL-----NIFL--SNHLKTNFBEKG---LIN 285
Oy      347  LGIKDRFVDIADTSLBGNDKKLATEAISBAKKEFEKTKTIPAEVYKRTLEIPEMDPKOP 406
Db      286  YGLKEGVGVFV-----RNPKMISF-----ELEKEIDNLSKTIINKELIYV----- 325
Oy      407  DKQESLDKLITD 419
Db      326  -SNKESYEKFLKE 337

```

Search completed: December 18, 2004, 01:27:31
Job time : 27.651 secs

```

Query Match      7.5%; Score 161; DB 2; Length 339;
Best Local Similarity 22.8%; Pred. No. 0.01;
Matches 85; Conservative 68; Mismatches 136; Indels 84; Gaps 18
59 GTVANDNSFNQSGMEAIQQQLGALTGTGEITSYVDST----ALEEGKXSSLANTKKWVWVLSG 114

```



```
QY 290 DTPDQSLVYTKKPKFFSILKNLGYVSFVSLDLYTKKNSRNLAGFEFGK---KQATV 345
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 311 DTPDQSKIFSPA-TVFETSIKHLRTIYQVLTDIWLKEDSKFLGSPRSFLTNPANATV 369
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 346 YLGIKDFVDIADTSLSGNDKKLATEAISEAKFEFEKTKTIPAEVRKTEIEPEM---P 402
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 370 YKGISDIFVGSNSTVADADKVKAKQEFLENTADFKQIQANPT-NYKSVGIPTMLIND 428
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 403 DKQDPKQOESL 413
|:::|
Db 429 NDADKNEKASL 439

RESULT 2
Q980L5 PRELIMINARY; PRT; 461 AA.
ID 0980L5
AC 0980L5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC TRANSPORTER XYLOSE-BINDING LIPOPROTEIN.
GN OrderedLocustNames=MYPV_3460;
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambud I., Heilig K., Wroblewski H., Vlati A., Kocha E.P.C.,
RA Moszer I., Dybvig K.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AI445564; CACI3519.1; -.
DR PIR; B90555; B90555.
DR MyPulst; MYPV_3460; -.
DR GO; GO:0008289; F:lipid binding; IEA.
DR InterPro; IPR003760; Bmp.
DR InterPro; IPR008107; Mycoplasma_p48.
DR InterPro; IPR011050; Pectin_lyase_like.
DR Pfam; PF02608; Bmp; 1.
DR PRINTS; PRO1733; LIPOPROTEIN48.
KM Complete proteome; lipoprotein.
SQ SEQUENCE 461 AA; 51096 MW; 6724D4D820809CE4 CRC64;

Query Match 25.6%; Score 552; DB 2; Length 461;
Best Local Similarity 31.2%; Pred. No. 3e-26;
Matches 133; Conservative 82; Mismatches 161; Indels 50; Gaps 11;

QY 13 ADONKOIT--DVKISGLVNER-----KSEIMAADKADKHFGLMAIYTAGTV 61
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 28 AQNNKTNNSLDSKRTIDLSQKEVTETQKVENKIKQASLETK-----VLTITAGSI 82
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 62 NDNSFNQSGWEALIQOLGALTGGEITSS-----VDSSTALEGKYSILANTNK 107
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 83 DDKSFNQOVYESQKTLDFDKAVKYSQNKAEANOHLKDNVINSVKKOLEQYKVALDRGY 142
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 108 NWVWLSGFQGDATATRLKLPENKQLFTEKNIIILIGIDMTDTEV-VIPTRGYINLYTKTE 166
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 143 TTYWLTFTFOQNELENLENLNDENNLRREKNEVKLIIGVWAPNANSKIPOGSLISLFRTE 202
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 167 EAGLAGYANASFLAKFPSPDPTKRSALVIGGGSIPAVTDFIAGYLGKIKAMN--LKNSD 224
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 203 EAGGQAGYASADFLGTGYRANNEAKRAISAFGGGPAVGTDFLNGFPESIRAMNSEAENAN 262
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 225 KKTITTDKLEINLGFVDQSTKERLEQIASKDKPSTLLAVAGPLTEIFSDII---ANQ 281
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 263 KKVYIVSENIVLDTGTF-IPNAEKNEVSVNVETGKSTISLPVAGPFGVVVDVLRKQTS 321
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 282 NDRYLIGVDDQSLVYTKKPKFFSILKNLGYVSFVSLDLYT-----KKSNSRNL 333
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 322 EDRFIVGVDDQSLSFNDSKRFFTSIVKNIAFPVQIILALITKBEESYILKENDKFL 381
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 334 AGFEFGKSAIVYIGIKDFVADIADTSLSEGNKKLATEAISEA--KKEFEKTKTIPAE 391
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 382 GS---NPKNLVLRKGLSAKVNITKRSVKESITQDQTSIQKALDKNANPNNSKKIKEM 438
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 392 VRKTE 397
|||
Db 439 TNGDLE 444

RESULT 3
Q6KIQ4 PRELIMINARY; PRT; 491 AA.
ID 06KIQ4
AC 06KIQ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Unspecified sugar ABC transporter binding protein.
GN OrderedLocustNames=NM080360;
OS Mycoplasma mobile.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=163K / ATCC 43663;
RA Birren B.W., Stange-Thomann N., Smith C., Decaprio D., Fisher S.,
RA Butler J., Calvo S., Elkin T., Fitzgerald M.G., Haefer N., Kodira C.D.,
RA Major J., Wang S., Wilkinson J., Nicol R., Nuebaum C.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017337; AAT27522.1; -.
DR InterPro; IPR008107; Mycoplasma_p48.
DR PRINTS; PRO1733; LIPOPROTEIN48.
KM Complete proteome.
SQ SEQUENCE 491 AA; 52817 MW; 2D1C707687771B41 CRC64;

Query Match 21.6%; Score 465.5; DB 2; Length 491;
Best Local Similarity 29.7%; Pred. No. 8.2e-21;
Matches 141; Conservative 80; Mismatches 165; Indels 89; Gaps 18;

QY 15 NOKKQITDVSKISGLVNERKSEIMAADKADKHFGLMAIYTAGTVNDNSFNQSGWEAL 74
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 32 NTGLRVTDNQFVDLVAAREAPATQRVANNSLPNSKTLIILITGAVVNDLSFNDSIIEAL 91
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 75 QOLGALTG--GEITSVDSSTA---ELEKGYSSILANTKNVWVLSGFQGDAPFRWLKIP 129
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 92 LEIGRQTKRPGNSFAETTAGTGDQLOQYDQALFFNHKFWVLITGFQGDAPFQWLDIGN 151
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 130 NKOLFTEKNIIILIGIDMTDENVIPTGRYINLYTKTEBACWLAGYANASFLAKKFPSPDPT 189
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 152 NRAEFIRKQVILVADITWTLLEVPQGFISIVRTQESSMIVGNVAKFISDHNHNN-- 209
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 190 KRAIVIGGGSIPAVNDFEIVAGIKAMWLKNS-----DKTKITT-DKIEY 236
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 210 -RTFNTFGGAPFVNTVNFAGFLQGLIDFN--NSTFLPEPETSITDCKULSFTPGDILNI 266
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 237 NLGFDVQDSTKERLEQIASKDKPST--LLAVAGPLTEIFSDIILANQD--RYLIGVDDQ 293
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 267 NTGFAV---TPRAATLQSIIVSGTQVFPVAGSLTTLVNSISQNSGQFVIGVDSQ 322
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 294 SLVYTKKPK--FTSILKNLGYVSFVSLDLY-----TKKSNS 330
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 323 AKAFSPDLATLFFSSVEKNVAGTTVALASLVIGTASTDPFNIITGSSSRFIPVTEKNS 382
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 331 RNLAGFEFGKSAIVYIGIR-----FVADIATSLSEGNKKLATEA---ISEAKKEF 380
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 383 SSLP-----LANADITKGFEVSTEPVDVGFSSALGKKTQSLVQANVGRSPAEVAD 437
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 381 -----EKKTK--TIPAEVRKTEIPEM--DKQDPKQOESLDKLITDIN 421
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 438 LAASLALFNKQKAKIATITP-----VQILPNSGGSGSTBEQILPNPLNELIKKIN 485
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```

DR InterPro: IPR003760; Bmp.
DR InterPro: IPR008107; Mycoplasma_p48.
DR Pfam: PF02608; Bmp; 1.
DR PRINTS: PR01733; LIPOPROTEIN48.
DR LipoProtein; Signal.
FT SIGNAL 1 22
SQ SEQUENCE 465 AA; 51149 MW; 60AD5448CFE03C66 CRC64;

Query Match 18.1%; Score 391; DB 2; Length 465;
Best Local Similarity 31.2%; Pred. No. 3,4e-16;
Matches 119; Conservative 76; Mismatches 128; Indels 58; Gaps 18;

QY 55 VTAGCTVNDNSFGNSGWEALQALGTGGETSVDSSTAELEK-----YSSL 102
DB 70 ITDGSVHDESFGNSGWEALVAKVSEGLDKAIV-SGNKILRNKYEPKQGLLEAVKNA 128
QY 103 ANTKNVMVLSGFQHGDAFTRLMKIPENK-QLPTEKNIIILGIMWT-----DTENV----- 152
DB 129 IDSGFRIVLCGFTHQASL---VGLDENYIKIKIDNNIIIFITVDPNLFETEDANVKTPIK 185
QY 153 -IPGRINLTYTEEGWLAGVANAFLAKKPSDPKSAIYIGGISPAVDTFAGY 211
DB 186 KIGGHLVPVIFDTKQAAVLAGALADYFSGVYKQDEPKRTIGAFGGIPWPAVSDPTAGT 245
QY 212 LAGIKANWLKSDKKTITTDKIEINLGFVDVDTSTKRLAQIAKQSPSTLLAVAGPL- 270
DB 246 FQGIIDMKHPEAKTSINETIELNTLF---TSGTQATTAISVYKATSYVAAGSL 302
QY 271 TELPSDI--IAMONRYLIGVDTQSLVYTKN-----KPTSLKMLGYSVSVSL 323
DB 303 TDTAKELIKLADK-DKFIIGVDADQ-----KNALKGHRIFTSVMKLIQAAVNNIADL 354
QY 324 YTKNSNRNL-AGFEFGKSKAT-VYLGIKD-----RPDIADTS-LEGNDKLAEALSEA 376
DB 355 YSKGENDLDLOPGEISKQKPTVFGYDTEKQYGVATSGLLDDKNDIANKALKDA 414
QY 377 KKEFEKTKTIPAEVRKTLIE 397
DB 415 TAYVQK-KT---EIQKSLK 430

RESULT 7
052311 PRELIMINARY; PRT; 428 AA.
ID 052311;
AC OS2311;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Membrane lipoprotein P48v.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
ON NCBI_TaxID=2115;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG18;
RA Submitted (Nov-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF036106; AAB99740.1;
DR GO: GO:0008289; F:lipid binding; IEA.
DR InterPro: IPR003760; Bmp.
DR InterPro: IPR008107; Mycoplasma_p48.
DR Pfam: PF02608; Bmp; 1.
DR PRINTS: PR01733; LIPOPROTEIN48.
DR LipoProtein; Signal.
SQ SEQUENCE 428 AA; 47862 MW; D9506B817E330EAA CRC64;

Query Match 17.8%; Score 384.5; DB 2; Length 428;
Best Local Similarity 27.0%; Pred. No. 7,8e-16;
Matches 116; Conservative 86; Mismatches 157; Indels 71; Gaps 17;

QY 31 NERKSEIIMAAKADANKH-----FGLNMAIVTAGCTVNDNSFGNSGWEAI 74

```

```

DB 27 NDNESNISFEKDISKTYTTNANGQVYKNAELLKLPVILITDEGKIDKSFNQSAREAL 86
QY 75 QQLGALIGEITVSSTAELEGYSLANTNKNVWVLSGFQHGDAFTRLMKIPENKOLF 134
DB 87 KAIKQGTIEINSVBS-SNFEASINLSAGHKIYNLWFKIQOSIKQY--IDAHBEEL 143
QY 135 TEKNIILIGIDMTDENVNPTGRYINLTYTEAGWLAGVANAFLAKKPSDPKSAI 194
DB 144 ERNQIKIIGIDF-DIETREYK--FVSLQFNIKESAPFTGYAISMWSEQ---DESKVIVA 197
QY 195 VIGGGSIPAVTDTFAGIAGIKAMNLKNSDKTKIT-TDKIEINLGFVDVDTSTKRL 253
DB 198 SFGGGAFFGYTTNEGPAKILLYNQGH--KSVKFTTSPVKKIDSGP-----TAGETKMT 250
QY 254 IAS-----KDRSTLLAVAGPLTEIFSDIANQNDRYLIGVDTQSLVYTKTKNK 303
DB 251 VINNVLSSTPADVKNPDHVLISVAGPAT--FETVKLANKQIVYIGVSDGAMT--QDKR 306
QY 304 FETSILKNLGYSVSVSLSDU-----YTKNSNRNLAGEFEFGKSAIVYLGIDRF 353
DB 307 ILTSLVKGHIQAAVETLDDILEKEGKPYVVKDKADKKWSHFTQ-----KEKW 358
QY 354 VDIADTSLBGNDKCL-ATEAISEAKKFEKTKTIPAEVRKTLIEPMD-KOPDQOE 411
DB 359 IGVAENHFSNTBEQAKINNKIKKAIKMKF-----LPEDFVKYINSRALKDGNKIDVSE 414
QY 412 SLQKLITDIN 421
DB 415 RLEAIIISAIN 424

RESULT 8
09RGK5 PRELIMINARY; PRT; 428 AA.
ID 09RGK5;
AC 09RGK5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Macrophage activating lipoprotein-404 precursor.
GN Name=malp; fermentans.
OS Mycoplasma fermentans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
ON NCBI_TaxID=2115;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SKS;
RX MEDLINE=99115554; Pubmed=9916088;
RA Calcutt M.J., Kim M.F., Kaipras A.B., Muhlradt P.F., Wise K.S.;
RT "Differential posttranslational processing confers intraspecies
RT variation of a major surface lipoprotein and a macrophage-activating
RT lipopeptide of Mycoplasma fermentans."
RL Infect. Immun. 67:760-771(1999).
DR EMBL: AF099211; AAD16395.1;
DR GO: GO:0008289; F:lipid binding; IEA.
DR InterPro: IPR003760; Bmp.
DR InterPro: IPR008107; Mycoplasma_p48.
DR Pfam: PF02608; Bmp; 1.
DR PRINTS: PR01733; LIPOPROTEIN48.
DR LipoProtein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 428 Potentlal.
SQ SEQUENCE 428 AA; 47835 MW; D03F0F47EAB21460 CRC64;

Query Match 17.7%; Score 382.5; DB 2; Length 428;
Best Local Similarity 27.0%; Pred. No. 1e-15;
Matches 116; Conservative 87; Mismatches 156; Indels 71; Gaps 17;

QY 31 NERKSEIIMAAKADANKH-----FGLNMAIVTAGCTVNDNSFGNSGWEAI 74
DB 27 NDNESNISFEKDISKTYTTNANGQVYKNAELLKLPVILITDEGKIDKSFNQSAREAL 86
QY 75 QQLGALIGEITVSSTAELEGYSLANTNKNVWVLSGFQHGDAFTRLMKIPENKOLF 134

```


Db 87 KAINKOTGIEINNEPS-SNFESAYNALSAGHKIWLNFGKHOQOQSIKQY--IDAAREEL 143
 QY 135 TEKNIITLIGDMTDETVIPTGRYINLTYTEAGWLAGYANASFLAKPPSPDPTKSAI 194
 Db 144 ERNOIKITIGIDF-DIEFEYK--FYSLOPNIKESAFPTGYAIAWSLEQ--DESKVVA 197
 QY 195 VIGGISPAVTDFTAGYIAGIKAMNLKNSDKTKI--TTDKIEINLGFVDOTSTKERLEQ 253
 Db 198 SPFGGAFPGVTTTFEGFAKILVYNQGH--KSKITYHTSPVKLDSGF-----TAGEMNT 250
 QY 254 IAS-----KDKSTLLAVAGPLTELFPSDIANQNDRIYLCVDTDSLVYTKTKNK 303
 Db 251 VINNVLSSTPADYKYNPHVILSVAGPAT--FETVRLANKQGYVIGVSDQGM1--QDKDR 306
 QY 304 FPTSILKNLGYVSFVLSDL-----YTKKSNRNLAGEFPGKKSATVYLGIDRF 353
 Db 307 ILTSVLKHIOAYETLIDLILEKEBGKRYVVKDKADKMSHFGQ-----KEKW 358
 QY 354 VDIADTSLGNDKKL-ATEAISEAKKEFEERTKTIPEAEVYKTLLEIPMPD-KQPKQOE 411
 Db 359 IGVAENHFSNTEBOAKINNKIKEMFKE---LPEDFVKYINSDKALDGNKIDNVSE 414
 QY 412 SLDKLTIDIN 421
 Db 415 RLEAIIISAIN 424

RESULT 9

ID 032417 PRELIMINARY; PRT; 428 AA.
 AC 032417;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE M61Ag.
 OS Mycoplasma fermentans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxId=2115;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=98022661; PubMed=9359703;
 RA Matsumoto M., Takeda J., Inoue N., Hara T., Hatanaka M., Takahashi K.,
 RA Nagasawa S., Akedo H., Seya T.;
 RT "A novel protein that participates in nonself discrimination of
 RT malignant cells by homologous complement.";
 RL Nat. Med. 3:1266-1270(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98241611; PubMed=9575196;
 RA Matsumoto M., Nishiguchi M., Kikkawa S., Nishimura H., Nagasawa S.,
 RA Seya T.;
 RT "Structural and functional properties of complement-activating protein
 RT M61Ag, a Mycoplasma fermentans gene product that induces cytokine
 RT production by human monocytes.";
 RL J. Biol. Chem. 273:12407-12414(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Seya T., Matsumoto M.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D64083; BAA23530.1; -
 DR EMBL; AB026157; BAA77211.2; -
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR InterPro; IPR003760; Bmp.
 DR InterPro; IPR008107; Mycoplasma_p48.
 DR Pfam; Pf02608; Bmp; 1.
 DR PRINTS; PR01733; LIIPROTEIN48.
 SQ SEQUENCE 428 AA; 47862 MW; F43B078F21ADDBE CRC64;

Query Match 17.6%; Score 379.5; DB 2; Length 428;
 Best Local Similarity 26.7%; Pred. No. 1.6e-15;
 Matches 115; Conservative 88; Mismatches 156; Indels 71; Gaps 17;

QY 31 NEKSEIMAKADANKH-----FGLNMAIVTAGVTVNDSFNQSGHEAI 74
 Db 27 NNDESINISFKKIDISKYTTTNANGQVKNABLLKLPVLTIDEGKIDDSFNQSAFEAL 86
 QY 75 QQLGALTGSEITVSDSTALEGKYSSTANTNKNVWLSGFQHDATRRMLKIPENQOLF 134
 Db 87 KAINKOTGIEINNEPS-SNFESAYNALSAGHKIWLNFGKHOQOQSIKQY--IDAAREEL 143
 QY 135 TEKNIITLIGDMTDETVIPTGRYINLTYTEAGWLAGYANASFLAKPPSPDPTKSAI 194
 Db 144 ERNOIKITIGIDF-DIEFEYK--FYSLOPNIKESAFPTGYAIAWSLEQ--DESKVVA 197
 QY 195 VIGGISPAVTDFTAGYIAGIKAMNLKNSDKTKI--TTDKIEINLGFVDOTSTKERLEQ 253
 Db 198 SPFGGAFPGVTTTFEGFAKILVYNQGH--KSKITYHTSPVKLDSGF-----TAGEMNT 250
 QY 254 IAS-----KDKSTLLAVAGPLTELFPSDIANQNDRIYLCVDTDSLVYTKTKNK 303
 Db 251 VINNVLSSTPADYKYNPHVILSVAGPAT--FETVRLANKQGYVIGVSDQGM1--QDKDR 306
 QY 304 FPTSILKNLGYVSFVLSDL-----YTKKSNRNLAGEFPGKKSATVYLGIDRF 353
 Db 307 ILTSVLKHIOAYETLIDLILEKEBGKRYVVKDKADKMSHFGQ-----KEKW 358
 QY 354 VDIADTSLGNDKKL-ATEAISEAKKEFEERTKTIPEAEVYKTLLEIPMPD-KQPKQOE 411
 Db 359 IGVAENHFSNTEBOAKINNKIKEMFKE---LPEDFVKYINSDKALDGNKIDNVSE 414
 QY 412 SLDKLTIDIN 421
 Db 415 RLEAIIISAIN 424

RESULT 10

ID 09RGX4 PRELIMINARY; PRT; 429 AA.
 AC 09RGX4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Macrophage activating lipoprotein-404 precursor.
 GN Name=malp;
 OS Mycoplasma fermentans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxId=2115;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT-2;
 RX MEDLINE=99115554; PubMed=9916088;
 RA Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;
 RT "Differential posttranslational processing confers intraspecies
 RT variation of a major surface lipoprotein and a macrophage-activating
 RT lipopeptide of Mycoplasma fermentans.";
 RL Infect. Immun. 67:760-771(1999).
 DR EMBL; AF09212; AAD16396.1; -
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR InterPro; IPR003760; Bmp.
 DR InterPro; IPR008107; Mycoplasma_p48.
 DR Pfam; Pf02608; Bmp; 1.
 DR PRINTS; PR01733; LIIPROTEIN48.
 KW Lipoprotein; Signal.
 FT SIGNAL 25
 FT CHAIN 1 429 Potential.
 SQ SEQUENCE 429 AA; 47933 MW; B6D08CF975AC3171 CRC64;

Query Match 17.6%; Score 379; DB 2; Length 429;
 Best Local Similarity 26.9%; Pred. No. 1.7e-15;
 Matches 116; Conservative 89; Mismatches 154; Indels 72; Gaps 18;

[illegible]

RESULT	11
Q9RGX6	PRELIMINARY; PRT; 429 AA.
AC	Q9RGX6;
DT	01-MAY-2000 (Tremblrel_13, Created)
DT	01-MAY-2000 (Tremblrel_13, Last sequence update)
DT	01-MAR-2004 (Tremblrel_26, Last annotation update)
DE	Macrophage activating lipoprotein-404 precursor.
GN	Name=malP;
OS	Mycoplasma fermentans.
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX	NCBI_TaxID=2115; [1]
RN	
RP	SEQUENCE FROM N.A.
RC	STRAIN=Incognitus;
RX	MEDLINE=99115554; PubMed=9916088; Calcult M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.; "Differential posttranslational processing confers intraspecies variation of a major surface lipoprotein and a macrophage-activating lipopeptide of Mycoplasma fermentans."; J Infect. Immun. 67:760-771(1999).
RL	Infect. Immun. 67:760-771(1999).
DR	EMBL; AF09210; AAD16394.1; -.
DR	GO; GO:0008289; F:lipid binding; IEA.
DR	InterPro; IPR003760; Bmp.
DR	InterPro; IPR008107; Mycoplasma_p48.
DR	Pfam; PF02608; Bmp; 1.
DR	PRINTS; PRO1733; LIPOPROTEIN48.
FT	Lipoprotein; Signal.
ET	CHAIN 1 24 Potential.
FT	CHAIN 25 429 macrophe activating lipoprotein-404.
QO	SEQUENCE 429 AA; 47961 MW; B6D09A8812AC3171 CRC64;

Query Match	17.6%	Score 379;	DB 2;	Length 429;
Best Local Similarly	27.0%	Pred. No. 1.7e-15;		
Matches 117; Conservative	88;	Mismatches 151;	Indels 78;	Gaps 18;

```
QY      3  NERKEIWAARADANKH-----FGLNMAIVYAGTGVNDNSPFQSGWEAL 74
      27  NNDSENISPEKEDJISKITTTNANGKQVYKNAELKIKPVLITDEGKIDDKSPFQSAFEL 86
Db
QY      75  QOLGALNGEITSVDNSTAELEKGYSSLANTNKNWVJLSGFHGDAFTMLKIPENKOLF 134
      87  KAINKOTIEINNPPS-SNPEASVNSALSAGHKIVLNGFGQOOSIKOY--IDAREEL 143
QY
```

```

QY      135  TEKNIITIGIDMTDITBENVIPTGRYINLYKTEBAGLAIANAASFLAKKFPSPDPTGRSAI 194
Db      144  ERNOIKIIGIDF-DIETEKW--FSLQPNIKESAFPTTQIALASWLSEQ---DESKRVA 197
QY      155  VIGGIGSPAYVDINGIYLAGIKAMUKNSDKKTKI-TUDKIEINIGFQDDNSTKRELEQ 253
Db      198  SFGGAGFPVYTFENEGFAPKGLIYNQOK--KSKKIYHTSPVKLDSG---PAGEKMT 250
QY      254  IAS-----KOKPSFLIYAVGPLETEI FSDI-IANONDRILIGVDPTDQSLVYKTKN 302
Db      251  VINNVLSSTPADVKYKNPHVILYSVAGPAT--FETVRLAANKGQYIVGSDQSMI--QDDK 306
QY      303  KEFTSILKNLGYSVESVLSDL-----YTKSNSHNLAGFEGKKSATVYIGIKOR 352
Db      307  RLTSVTLKHIKQAVYETLLDLLEKEBGYKPYVVDKADKKMSHGTQ-----KEK 358
QY      353  FVDIADTSLLEGNDKYL-ATEAISEAKKEF---EKKTIIPAEVYKTLIEIPMPDKOP 407
Db      359  WIGVAENHFNSMTEQAKINNKIKAEIKXEFKELPEDFVKYINSDKVLK-----DGNKID 411
QY      408  KQOESLDKLITDIN 421
Db      412  NVSERLEAIIISAIN 425

```

RESULT	12			
ID	09RGX7	PRELIMINARY;	PRT;	428 AA.
AC	09RGX7;			
DT	01-MAY-2000	(TEMBLrel. 13, Created)		
DT	01-MAY-2000	(TEMBLrel. 13, Last sequence update)		
DT	01-MAR-2004	(TEMBLrel. 23, Last annotation update)		
DE	Macrophage activating lipoprotein-404 precursor.			
GN	Name=malD; fermentans.			
OS	Mycoplasma.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_TaxID=2115;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=IT-29/1;			
RA	MEDLINE=91115514; PubMed=9916088;			
RX	Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;			
RT	"Differential posttranslational processing confers intraspecies			
RT	variation of a major surface lipoprotein and a macrophage-activating			
RT	lipoprotein of Mycoplasma fermentans.";			
RL	Infect. Immun. 67:760-771(1999).			
DR	EMBL; AF099209; AAD16393.1; --			
DR	GO; GO:0008289; F:lipid binding; IEA.			
DR	InterPro; IPR003760; Bmp.			
DR	InterPro; IPR008107; Mycoplasma_p48.			
DR	Pfam; PF02608; Bmp; 1.			
DR	PRINTS; PR01733; LIPOPROTEIN48.			
KW	LIPOPROTEIN; Signal.			
FT	SIGNAL	1	24	Potential.
FT	CHAIN	25	428	macrophage activating lipoprotein-404.
SQ	SEQUENCE	428 AA;	47960 MW;	DB569C0BA969CFE1 CRC64;

Query Match	17.5%	Score 377.5;	DB 2;	Length 428;
Best Local Similarity	27.8%	Pred. No. 2.1e-15;		
Matches 110; Conservative	84;	Mismatches 146;	Indels 55;	Gaps 16;

[illegible]

QY	230	TTDKIEINLPGDVDTSTKKELEQIAS-----KDKSTLLAAVGLTEIFSDI	278
DB	221	YHTSVKLDSGF-----TAGKMTVINNVLSSTPADVKYNPHVLSVAGPAT--FETVR	283
QY	279	ANONDRYLIGVDTOQLSVYTTKTKNKFSTILKNLGSVSVLSDU-----YTKKS	328
DB	284	LANKQYVIGVDSDDGMI--QDKRILTSVLKHIKQAVYETLLDLILEKEBEGKRYVVD	341
QY	329	NSRNLAGEPEFKKATVVLGIKDRVVDIADTSLSEGNKKL-ATEAISLAKKEFEKTKTI	387
DB	342	KKAKKWSHFSTQ-----KEKIIQVAENHFSNTEQDAKINNKIKAIKMFKE-----L	389
QY	388	PAEEYKTLSEIPEMPD-KOPDKQOSRLDKLITDIN	421
DB	390	PEDFVKYINSDKALDKGNKIDNVSERLAIIISAIN	424
RESULT 13			
Q9RGX3	ID	PRELIMINARY;	PRT; 428 AA.
AC	Q9RGX3		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)	
DE	Macrophage activating lipoprotein-404 precursor.		
GN	Name=malp:		
OC	Mycoplasma fermentans.		
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TaxID=2115;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K7;		
RX	MEDLINE=99115554; PubMed=9916088;		
RA	Calcutt M.J., Kim M.F., Kaipaa A.B., Muhlradt P.F., Wise K.S.;		
RT	"Differential posttranslational processing confers intraspecies		
RT	variation of a major surface lipoprotein and a macrophage-activating		
RT	lipopeptide of Mycoplasma fermentans.";		
RL	Infect. Immun. 67:760-771(1999).		
DR	EMBL: AF099213; AADI6397.1: -		
DR	GO: GO:0008289; F:lipid binding; IEA.		
DR	InterPro: IPR003760; Bmp.		
DR	InterPro: IPR008107; Mycoplasma_p48.		
DR	Pfam: PF02608; Bmp; 1.		
DR	PRINTS; PR01733; LIPOPROTEIN48.		
KM	Lipoprotein; signal.		
FT	CHAIN	1 24	Potential.
FT	CHAIN	25 428	macrophage activating lipoprotein-404.
SO	SEQUENCE	428 AA; 47875 MW; C7A536B409A60132 CRC64;	
Query Match 17.4%; Score 374.5; DB 2; Length 428;			
Best Local Similarity 26.7%; Pred. No. 3,3e-15;			
Matches 115; Conservative 86; Mismatches 158; Indels 71; Gaps 17			
QY	31	NERKSEIQAARADANKH-----FGLNMAIVTAGGTVDNDSFNQSGWEAI	74
DB	27	NNDESINIFPEKEDISKYTTTANGKQVVAELKLKPLVLTDESKIDDKSFNQGAFAEL	86
QY	75	QQLGALTGEITSVDSSTAELEGKSSSLANTKNVWVLSGPHGADFTRMILKIPENKOLF	134
DB	87	KAINKQTGTEINSVSPS-SNFSASNAISLSAGHKIWLNGFPHGQSIQY--IDAHEEL	143
QY	135	TEKNIILIGIMTDEENVIPTGRYVNLVYKTEAGMLAGYANASEFLAKKFPSPKRSAL	194
DB	144	ERNQIKIGIDF-DIETEKYK--FYSLQENIKERAPFTTQYALASWLSSEQ--DESKRYVA	197
QY	195	VIGGISIPAVYDFIAGYLAGIKAMNKLKSDKTKY--TDKIEINLPGDVDTSTYERLEQ	253
DB	198	SFGGGAFFGVGVTTFNGCFAGILYVQKK--KSSKIYHTSPVXLDGSGF-----TAGKMT	250
QY	254	IAS-----KDKPSTLLAAVAPLREIFSDIANONDRYLIGVDTOQLSVYTTKTKN	303
DB	251	VINNVLSSTPADVKYNPHVLSVAGPAT--FETVRANKQYVIGVDSDDGMI--QDKOR	306

QY	304	FFTSILKRLKGVSVESVLSDL-----YTKKSNBRNLGFEFGKKSATYVLAGIKDRF	353
DB	307	ILTSVLKRIKQAVYETLDDLILKEEGYKPYVVDKKADKKSHSTGTQ-----KEKM	358
QY	354	VDIADTSLBGNDKKL-ATEAISAEKKKEFEKTKTIPAEVVKTLTEIPENP-KQDPKQOE	411
DB	359	IGVANEHPSNTEEGOKKINKKIKKAIKQWKE-----LPDEFVKYINSDKALKDQGNKIDNVSE	414
QY	412	SLDKLTIDIN 421	
DB	415	RLEALISAIN 424	
RESULT 14			
QY	Q9R3N6	PRELIMINARY; PRT; 428 AA.	
AC	Q9R3N6;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)		
DE	Lipoprotein MALP-404 precursor (Macrophage activating lipoprotein-404 precursor).		
GN	Name=malp; Synonyms=malp;		
OS	Mycoplasma fermentans.		
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TaxId=2115;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PG18, M39A, and M70B;		
RX	MEDLINE=99115554; PubMed=99116088;		
RA	Calcutt M.J., Kim M.F., Karpas A.B., Muhlradt P.F., Wise K.S.;		
RT	"Differential posttranslational processing confers intraspecies		
RT	variation of a major surface lipoprotein and a macrophage-activating		
RT	lipopeptide of Mycoplasma fermentans."/		
RL	Infect. Immun. 67:760-771(1999).		
DR	EMBL: AF100324; AAD25736.1; -		
DR	EMBL: AF099214; AAD16398.1; -		
DR	EMBL: AF099215; AAD16399.1; -		
DR	GO; GO:0008289; F:lipid binding; IEA.		
DR	InterPro; IPR003760; Bmp.		
DR	InterPro; IPR008107; Mycoplasma_p48.		
DR	Pfam; PF02608; Bmp; 1.		
DR	PRINTS; PR01733; Lipoprotein48.		
KW	Lipoprotein; Signal.		
KW	Lipoprotein; Signal.		
FT	SIGNAL	24	Potential.
FT	CHAIN	1 25 38	macrophage activating lipopeptide MALP-2.
SO	SEQUENCE	428 AA; 47863 MW; 59808324E218F03C CRC64;	
QY	Query Match	17.3%; Score 373.5; DB 2; Length 428;	
	Best Local Similarity	26.7%; Pred. No. 3.8e-15;	
	Matches 115; Conservativity 87; Mismatches 157; Indels 71; Gaps 17		
QY	31	NERKSEIMAAVADANKH-----FGANMAIVTAGTVNDNSFNSQGWEMAI	74
DB	27	NNDESNIISFKEKDISKTYTTNANGQVYKAMLELTKLKYLTIDEGKIDDKSNQSAFEL	86
QY	75	QOLALGTGEITSYDSTAELEGKYSISLANTKNVAVLSGFQHGDAFTRMKIPENKOLF	134
DB	87	KAINMQGTIEINSVPS-SNPEASANSMALSGHKIKVWLGFGHQGSIKQY--IDAREEL	143
QY	135	TEKNITIIIGIDMTDTENVIPTRGYINLTLYKTEEAGMLAGYANASLAKKFPSPDPTKSAI	194
DB	144	ERNOIKIIGIDF-DIETEKW--FYSLOFNIKESAFITGYALASWLSEQ---DESKRVA	197
QY	195	VIGGISPAVVDFTLAGYLAGIKANMLKNSDKKTKI-TTDKIEINLGPVDQDSTTERLEQ	253
DB	198	SPGVGAFPGVTFTEGFAKGLIYNQKQK--KSSKYHTSPVYLDGSE---TAEGKANT	250
QY	254	IAS-----KDRPSTLLAVAGPLTEIFSDITIANQNDRYLIGVDTPDOSLVYTKTKXK	303
DB	251	VINNVLSSTPADVKKYNNPHVILISVAGPAI--FETVRLANKGVIVIGVDDQGMH--QDQDR	306
QY	304	FFTSILKRLKGVSVESVLSDL-----YTKKSNBRNLGFEFGKKSATYVLAGIKDRF	353

[illegible]

RESULT 15

ID	Q8RCH2	PRELIMINARY;	PRT;	352 AA.
AC	Q8RCH2			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Surface lipoprotein.			
GN	Name=Med; OrderedLocustNames=TFE0457;			
OS	Thermomicrobacter tengcongensis.			
OC	Bacteria; Firmicutes; Clostridia; Thermomicrobacteriales;			
OC	Thermomicrobacteriaceae; Thermomicrobacter.			
OX	NCBI_TaxID=119072;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=MB4;			
RX	MEDLINE=121992816; PubMed=11997336;			
RA	Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,			
RA	Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,			
RA	Tan H., Chen R., Wang J., Yu J., Yang H.,			
RT	"A complete sequence of the T. tengcongensis genome.",			
RL	Genome Res. 12:689-700(2002).			
DR	EMBL; AEO13017; AAM23740.1; -			
DR	GO; GO:0008289; F:lipid binding; IEA.			
DR	InterPro; IPR003760; Bmp.			
DR	Pfam; PF02608; Bmp; 1			
KM	Complete proteome; Lipoprotein.			
QO	SEQUENCE 352 AA; 38052 MW; DCA7990EA271A23B CRC64;			

Search completed: December 18, 2004, 01:26:33
Job time : 121.543 secs